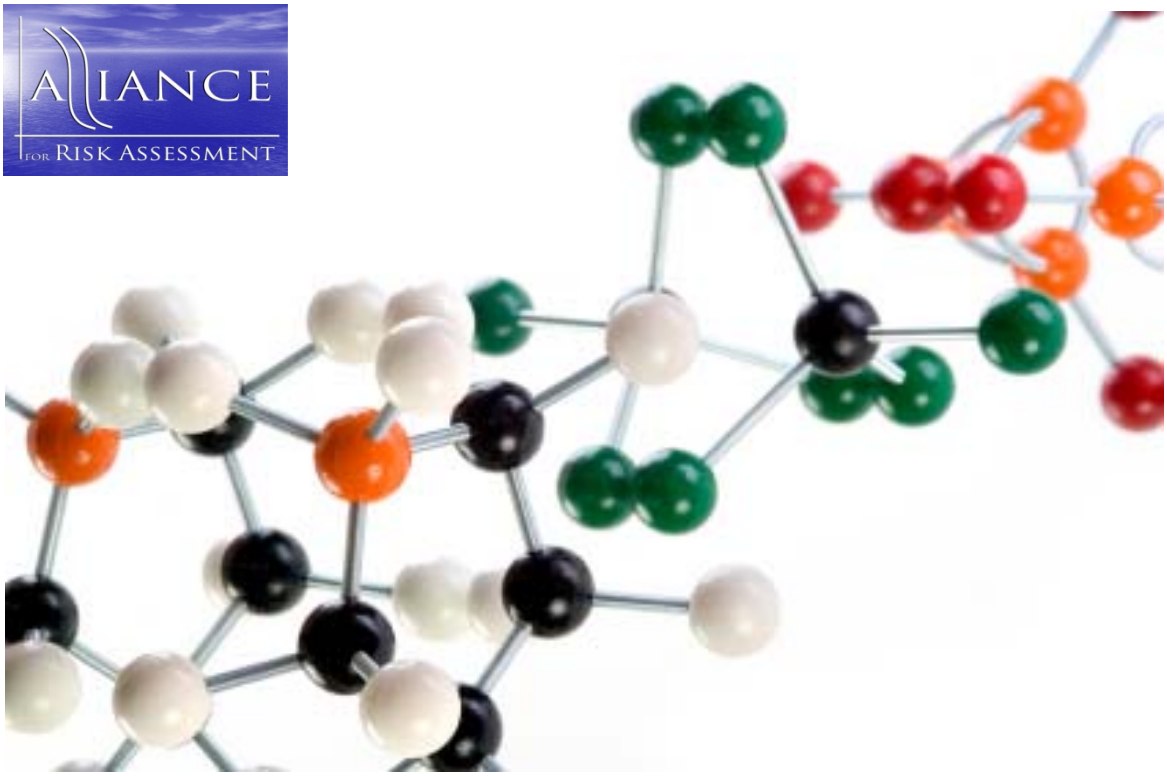


Beyond Science and Decisions:
From Problem Formulation to Dose-Response
Report from Workshop VI

Workshop Held:
May 28-30, 2013
Arlington, VA
Potomac Yard, U.S. Environmental Protection Agency

Report Prepared By:
Toxicology Excellence for Risk Assessment (TERA)
Lynne Haber (haber@tera.org)

October 10, 2013

Table of Contents

Introduction

Workshop Scope and Objectives

The workshop series, Beyond Science and Decisions: From Problem Formulation to Dose-Response continues and expands upon the discussion initiated by the National Academy of Science report: *Science and Decisions: Advancement of Risk Assessment* (NRC, 2009). The workshops utilize a multi-stakeholder format to support the development of a practical and solution-oriented compendium of risk assessment methods. Conducted under the aegis of the Alliance for Risk Assessment (*ARA*), the workshop series explores both currently available and evolving methodologies, through the development and application of case studies. The workshop series is based on the fundamental premise that the appropriate methodologies for dose-response assessment need to be based on objectives specific to the intended application; this will include varying levels of analysis.

The workshop series continues to advance the framework of *ARA* (2012) on problem formulation and dose-response analysis (beta version available at http://chemicalriskassessment.org/methods/).

The purpose of this workshop report is to document and communicate the workshop results to the workshop participants and interested others. The report contains summaries of the Science Panel discussions with the authors of invited presentations, as well as the Science Panel review of case studies presented at the workshop. The draft Workshop report was reviewed by the panel and presenters, and their comments have been incorporated into the final report.

Science Panel

The standing Science Panel chosen by the *ARA* Steering Committee prior to Workshop IV continued its service for Workshop VI. Panel biographies are provided in Appendix 1, as well as at http://www.allianceforrisk.org/Workshop/Panel.htm. The Science Panel for Workshop VI consisted of the following, including standing panel members and one *ad hoc* member:

- *Richard Beauchamp, Texas Department of State Health Services*
- *James S. Bus, The Dow Chemical Company*
- *Rory Conolly, U.S EPA National Health and Environmental Effects Research Laboratory*
- *Michael L. Dourson, Toxicology Excellence for Risk Assessment*
- *Annie M. Jarabek, U.S. EPA, Office of Research and Development (ad hoc)*
- *R. Jeffrey Lewis, ExxonMobil Biomedical Sciences, Inc.*
- *Bette Meek, McLaughlin Centre for Population Health Risk Assessment, University of Ottawa*
- *Gregory Paoli, Risk Sciences International*[1]

[1] Member of the NAS *Science & Decisions* panel

- *Alan Stern, New Jersey Dept of Environmental Protection*

Workshop VI Organization

The workshop was organized by the Dose-Response Advisory Committee (DRAC) on behalf of the more than 50 workshop sponsors. The DRAC determined the agenda (see Appendix 2) in consultation with the Science Panel. The sponsors of the workshop series are listed at http://www.allianceforrisk.org/ARA_Dose-Response_Sponsors.htm. Additional support for this workshop was provided by the U.S. Environmental Protection Agency (US EPA), who hosted the workshop. The workshop included both invited presentation on topics of interest to and requested by the Science Panel, and case studies being reviewed by the Science Panel. The workshop was open to the public for both in-person participation and participation via webcast. Public comments were invited at selected times during the workshop. The list of workshop participants is included in Appendix 3 of this report.

The following were invited presentations at the meeting. Summaries of the panel discussions following the presentations are provided in this report.

- Ken Olden, U.S. EPA, National Center for Environmental Assessment. Plenary address.
- Thomas Hartung, John Hopkins Bloomberg School of Public Health, Centers for Alternatives to Animal Testing. *Pathway-Based Regulatory Toxicology and Alternatives to Animal Testing.*
- Bette Meek, University of Ottawa. *International Developments on Mode of Action.*
- Tim Pastoor, Syngenta Crop Protection, Inc. *The HESI RISK21 Program.*
- Ted Simon, Ted Simon, LLC. *The HESI RISK21 Quantitative Key Events Dose Response Framework (Q-KEDRF).*

Much of the workshop was dedicated to review of case studies. Each review began with a presentation by the case study author(s) on key elements, followed by a panel discussion. The purpose of the panel discussion was to identify areas for additional development of case studies and/or refinement of methods. The following case studies were presented:

- Robinan Gentry, Environ International Corporation; Tom Starr, TBS Associates; Jim Swenberg, University of North Carolina Chapel Hill; Jeffry Schroeter, Applied Research Associates. *Endogenous Formation Implications for Formaldehyde Carcinogenicity.*
- Lorenz Rhomberg, Gradient; Lisa Bailey, Gradient. *Case Study: Hypothesis-Driven Weight of Evidence Review for Naphthalene Carcinogenicity.*
- Roberta Grant; Joseph "Kip" Haney; Allison Jenkins; Texas Commission on Environmental Quality (TCEQ). *Interpretation of 24-hour Sampling Data*. *Case Study A: Texas Commission on Environmental Quality Approach.*
- Denis Jugloff; Julie Schroeder; Ontario Ministry of Environment (MOE). *Interpretation of 24-hour Sampling Data*. *Case Study B: Ontario Ministry of the Environment Approach.*

Except for Dr. Olden's talk, which was not accompanied by slides, all presentations are available at http://www.allianceforrisk.org/Workshop/WS6/WS6casestudies.html. Similarly, with the exception of the keynote talk, the abstracts for all invited talks were provided by the speakers.

Panel Discussions of Presentations

Keynote Talk, Dr. Kenneth Olden[2]

Dr. Olden presented his vision for EPA's National Center for Environmental Assessment (NCEA) and steps being taken to achieve that vision with respect to improving the Integrated Risk Information System (IRIS). He stated that NCEA is seeking to enhance best practices, and soliciting ideas from across the community for improving the IRIS program and its process for providing toxicological reviews and assessments. His vision is for NCEA to be a world-class organization, providing first-rate information in which federal agencies, other government agencies, and the American people can have confidence. In order to achieve this vision for the IRIS program, NCEA is striving to address the following enhancements to the IRIS program:

1. The IRIS process and the toxicity review documents are changing. The IRIS program is striving to enhance its assessment process to improve quality, transparency, and productivity
2. Achieve consistency in peer review
3. A transparent, consistent, concise, and easy to understand approach is needed.
4. Better communication among risk assessors, risk managers, and stakeholders is needed; risk managers should be involved in the planning and scoping up front.
5. Collegiality among scientists, and government and non-governmental stakeholders should be promoted, with relationships based on trust and respect.
6. A better plan for the use of the work force is needed.
7. The IRIS program needs to be proactive and address scientific challenges encountered in applying new science and methods. Public debate aids in scoping the issues and fostering trust.

Initial enhancements to the IRIS program have focused on process, and Dr. Olden noted that a revised process would be released in early summer, and that public debate on the proposal would be welcomed. The IRIS process announcement is available at http://www.epa.gov/iris/process.htm. The second part of improvements to the IRIS program will address advancements to the science, including the incorporation of new methods such as omics, informatics, and predictive modeling. Engagement of the scientific and stakeholder communities will be important, as well as case studies to illustrate and test new approaches. Dr. Olden noted the environmental health roundtable related to toxicogenomics when he was at NIEHS as a model for bringing together stakeholders from a variety of perspectives.

Dr. Olden also noted a need to broaden the definition of environment to include other stressors, such as physical agents, social and behavioral factors, nutrition, stress, and exercise. He also

[2] The rapporteur's summary is provided, since no abstract was provided for Dr. Olden's talk and Dr. Olden did not use slides.

suggested the use of biomarkers of effect to evaluate cumulative exposures in a community setting.

In summary, Dr. Olden called upon the scientific and stakeholder communities to work together to move the field forward, and to work together to build trust and respect in addressing environmental health issues, recognizing that all are committed to building a better world.

DISCUSSION:

In response to a panelist question about how to link new technology with what is known about world of exposure, Dr. Olden discussed designing a neighborhood study measuring key exposures (air, water, stress, economic factors, crime, etc.) and then comparing neighborhoods with large commonalities or differences in health outcomes. The cumulative exposure assessment could be replicated with animals under controlled conditions for the endpoint (biomarker) of interest. This leads to well coordinated and planned series of studies that are more informative than multiple isolated small studies.

Noting the difficulty of funding multidisciplinary efforts, a panelist inquired about the opportunities for developing a niche that funds risk assessment. Dr. Olden responded that progress has been made in this area, but there is much further to go. One approach would be to lobby National Institutes of Health (NIH) directors to consider different scientific paradigms, or to encourage different institutes to partner.

With regard to the idea of addressing conflict, Dr. Olden noted that the idea that one of the panelists has raised of building a risk assessment community involving all stakeholders is a good one. The NIEHS environmental health roundtable included well-recognized and experienced scientists from universities, industries, environmental groups, EPA, ATSDR, etc. Such groups can be productive even if they only generate ideas; others can then obtain the resources to work on those ideas.

In response to a panelist question about strategies for encouraging change in a timely fashion, Dr. Olden stated that it is important to get buy-in and get people excited about the potential outcomes. For the changes at NCEA and in the IRIS program, the approach has included recognizing how the proposed changes would benefit the Agency and the nation, acknowledging the public service ideals that led most EPA employees to work at the Agency, honestly engaging people in asking for their input, and building relationships and enthusiasm for new ideas.

In response to a panel member question, Dr. Olden clarified that the idea of broadening the definition of environment does include supporting sustainability. A member of the audience recommended that social factors should be considered separately from the environmental exposures that are subject to regulation and that have been the traditional focus of IRIS. Dr. Olden replied by noting that disease is caused by an interaction of genetics, environment, and social behavior that get imprinted on the epigenome. He recommended that this interaction be considered based on factors turning gene expression on and off.

Another audience member brought up the challenges of evaluating the weight of evidence (WOE) in doing assessment, such how much data are sufficient to express risk and how to consider the results of a study that has not been replicated. Dr. Olden responded that there is rarely one study that changes the paradigm (for evaluating a chemical); instead, each study provides a piece of the overall understanding. It is important to fully use the available data and really think about the data. He noted that decisions need to be made, and it is not possible to wait until the science is fully settled. As an example, he noted the work of the scientific advisory committee for IRIS assessments, which ensures that many perspectives have been received and considered by the time an assessment is published. This committee shows that it is possible for 25 people to come to consensus that a certain approach is the best possible at the current time.

Recent International Developments in Mode of Action/Adverse Outcome Pathway Analysis, Dr. Bette Meek

ABSTRACT:

The WHO/IPCS mode of action/human relevance (MOA/HR) framework has recently been updated (see Figure 1) to reflect evolving experience in its application and to incorporate recent developments in toxicity testing and non-testing methods. The modified framework is incorporated within an iterative roadmap, encouraging continuous refinement of problem formulation, mode of action based testing strategies and risk assessment. It can be used as originally intended, where the outcome of chemical exposure is known, or in hypothesizing potential effects resulting from exposure, based on information on putative key events in established modes of action from appropriate *in vitro* or *in silico* systems and other evidence.

The implications of the experience acquired in application of the framework in addressing documented (adverse) effects to inform the more limited knowledge base in these more predictive applications are addressed. This is illustrated in various case examples including the use of mode of action analysis in prioritizing substances for further testing, in guiding development of more efficient testing strategies and in identifying critical data gaps and testing strategies in read-across.

In addition to clarifying terminology related to the essentially conceptually synonymous terms of mode of action and adverse outcome pathways, the Bradford Hill (BH) considerations have also been articulated as a basis to simplify their application in considering weight of evidence for hypothesized modes of action.

Contribution of these developments to international initiatives on advancement of integrated test strategies based on evolving methods were also addressed.

DISCUSSION:

In response to a panelist question about the 2013 workshop, Dr. Meek clarified that the publication has been submitted and that the workshop report will be posted soon on the European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC) site (see http://www.ecetoc.org/). The MOA wiki noted in her presentation will be on the EPA website,

and is being developed in collaboration with the Joint Research Centre of the European Commission (JRC) and the Army Corp of Engineers. Another panel member noted the importance of toxicokinetics in determining the shape of the dose-response curve and expressed concern that the revised MOA framework does not explicitly incorporate any consideration of the impact of toxicokinetics (TK) on subsequent toxicodynamics (TD) and that both TK and TD define an MOA. Dr. Meek agreed with the importance of considering toxicokinetics and noted that it is considered more in the context of evaluating interspecies concordance addressed subsequently in the MOA framework, rather than in assessing weight of evidence for hypothesized MOAs. She also clarified that absorption is not generally considered a key event in mode of action in this framework, though some MOA analyses have been based on this premise. Another panelist noted the importance of toxicokinetics in advancing *in vitro* methods and that dosimetry is key for linking exposure to effect. In response to a panelist comment, Dr. Meek agreed that MOAs and adverse outcome pathways (AOPs) are the same, but noted that evaluation of human exposure is not part of the MOA evaluation.

A panel member agreed with the comments in Dr. Meek's presentation about focusing on confidence rather than uncertainty, with more data increasing the level of confidence, but asked for clarification about "fit for purpose". Dr. Meek replied that a key element of "fit for purpose" is recognizing the drivers of the assessment, particularly legislative drivers. Some legislative requirements require greater predictivity up-front, and a focus on MOA aids in such predictivity. One also needs to consider the goal of the weight of evidence evaluation. For example, review of an entire dataset is not needed if the goal is "read across" from other chemicals for prioritization. The goal of the international initiatives described in the presentation is to provide some examples and to encourage people to begin to apply more predictive approaches that integrate advancements in technology.

In response to a panelist comment and clarification agreeing with the statement that it is incorrect to describe a chemical as having multiple MOAs, Dr. Meek reiterated that, based on her discussions with the biological community, the proper way to describe such a situation is one MOA. The MOA entails several key events, each of which may result from different (sometimes) competing mechanisms and/or pathways, although these converge at a late stage to produce the (adverse) effect. It is important to work with the research community to assimilate all of the information in order to describe one integrated MOA. As an example, Dr. Meek clarified that if a chemical is a mutagen and also stimulates clonal replication, then the MOA involved both of these as key events.

Pathway-Based Regulatory Toxicology and Alternatives to Animal Testing, Dr. Thomas Hartung

ABSTRACT:
A mechanistic toxicology has evolved over the last decades, which is effectively relying to a large extent on methodologies which substitute or complement traditional animal tests. The biotechnology and informatics revolution of the last decades has made such technologies broadly available and useful. Regulatory toxicology has only slowly begun to embrace these new approaches. Major validation efforts, however, have delivered the evidence that new approaches do not necessarily lower safety standards and can be integrated into regulatory safety

assessments, especially in integrated testing strategies. Political pressures especially in the EU, such as the REACH legislation and the 7th amendment to the cosmetic legislation, further prompt the need of new approaches. In the US, especially the NAS vision report for a toxicology in the 21st century and its most recent adaptation by EPA for their toxicity testing strategy have initiated a debate how to create a novel approach based on human cell cultures, lower species, high-throughput testing and modeling. The report suggests moving away from traditional (animal) testing to modern technologies based on pathways of toxicity. These pathways of toxicity could be modeled in relatively simple cell tests, which can be run by robots. The goal is to develop a public database for such pathways, the Human Toxome, to enable scientific collaboration and exchange.

The problem is that the respective science is only emerging. What will be needed is the Human Toxome as the comprehensive pathway list, an annotation of cell types, species, toxicant classes and hazards to these pathways, an integration of information in systems toxicology approaches, the in-vitro-in-vivo-extrapolation by reversed dosimetry and finally making sense of the data, most probably in a probabilistic way. The NIH is funding our Human Toxome project (http://humantoxome.com) since September 2011 by a transformative research grant. The project involves US EPA ToxCast, the Hamner Institute, Agilent and several members of the Tox-21c panel. The new approach is shaped around pro-estrogenic endocrine disruption as a test case.

Early on, the need for quality assurance for the new approaches as a sparring partner for their development and implementation has been noted. The Evidence-based Toxicology Collaboration (EBTC, http://www.ebtox.com) was created in the US and Europe in 2011 and 2012, respectively. This collaboration of representatives from all stakeholder groups aims to develop tools of Evidence-based Medicine for toxicology, with the secretariat run by CAAT. All together, Tox-21c and its implementation activities including the Human Toxome and the EBTC promise a credible approach to revamp regulatory toxicology.

DISCUSSION:

In response to a panelist question about how the concepts of evidence-based toxicology (EBT) might contribute to guidance that is being developed for evaluating weight of evidence (WOE), Dr. Hartung noted the commonalities between the EBT and WOE initiatives, with the difference that EBT is a formalistic approach, based on the Cochrane method, meta-analysis, and other quantitative tools, while the WOE methods focus on pragmatic approaches. The development of quality measures, such as systematic Klimisch scores for *in vitro* and *in vivo* studies can aid in combining data across different study types.

With regard to the software tools noted in the presentation as being developed under a National Institutes of Health (NIH) grant, Dr. Hartung noted that the goal is to make them publically available as part of a high quality toolbox. A panel member praised the overall initiative, and noted that other initiatives have highlighted the importance of having such tools available in an open source form.

In response to a panelist question about how an "organ on a chip" would capture interactions between organs (e.g., a metabolite is formed in the liver, and then is circulated in the blood stream and causes effects in other organs), Dr. Hartung agreed that such applications currently fall into his "pie in the sky" category. While the ultimate goal is to be able to capture such interactions, substantial development is needed before this goal can be achieved. The current initiatives have a strong emphasis on toxicokinetics and microfluidics, in order to be able to start to address such issues.

In response to a question about the role that Bayesian analysis and probabilistic networks play in the projects, Dr. Hartung noted that the bioinformatics subgroup is using Bayesian analysis to address issues such as determining which toxicity test should be done next (based on the currently available information) to provide the most additional information. Dr. Hartung recommended that risk assessment should move towards thinking in terms of a probability of a certain hazard being manifested, rather than yes/no answers for cancer and other endpoints. Probabilistic techniques can be used to integrate the data across many different studies. A workshop on this topic is planned with the possible development of a steering group. Some test cases are under development, such as application to test methods for skin sensitization tests, and larger projects with the fragrance industry.

The HESI RISK21 Program, Dr. Tim Pastoor

ABSTRACT

The RISK21 Roadmap is a straightforward, efficient, and systematic way to achieving a transparent assessment of human health risk to chemicals. The Roadmap is a problem-formulation based, exposure-driven, and tiered methodology that seeks to derive only as much data as is necessary to make a safety decision. This presentation usea a RISK21 "pseudomethrin" case study to focus on how existing information and tiered data development can be used in human safety decision making. Pseudomethrin is envisioned as the next pyrethroid in a group of 11 well-tested pyrethroids. The problem formulation asks how much data is needed to decide that pseudomethrin could be used on bed netting with reasonable certainty of no harm. The RISK21 exposure-driven process that utilizes prior information on existing pyrethroids shows what toxicological data would be necessary to achieve this decision.

DISCUSSION

Panel members discussed how the Risk21 approach compares with existing tiering systems, and can build on previous related work. It was noted that there is potential for harmonizing the Risk21 approach with other tiering approaches, but differences among approaches can lead to confusion. For example, the Risk21 approach does not include the less data-informed approaches that others have used for evaluating exposure in Tier 0. Similarly, a panel member recommended that MOA be included at lower tiers, not only at higher tiers. Dr. Pastoor replied that the Risk21 team did not overtly build on other approaches. Instead, they used common sense, starting with something simple, adding additional complexities, and then condensing the approach to a visual graphic. With regard to MOA, Dr. Pastoor noted that the (data-poor)

pyrethroid case study could use *in vitro* data based on an understanding of MOA. The panel member recommended that the diagram be modified to reflect the potential for such approaches to be mutually informative and to include MOA information at lower tiers.

Dr. Pastoor noted that Risk21 includes participants from 12 countries, and there are ongoing outreach efforts to encourage uptake of the approach. The results of the Risk21 efforts will be posted on the web at risk21.org, which will have an automated system facilitating the use of the roadmap and matrix. Overall, the Risk21 team is very interested in transparency and accountability. The panel member offered to share additional ideas on engagement strategies to integrate the Risk21 work with work from other groups.

Another panel member asked for clarification of the amount of information needed to make decisions, noting that staying away from the dotted line in the diagram (see Figure 2) that defines exposure/toxicity combinations that may be of concern is itself an implied decision. The panel member noted that different criteria maybe be used for different questions – e.g., comparing technologies vs. is 'X' safe enough, vs. a risk-risk comparison, such as the risk of a chemical exposure compared to the risk from malaria. Dr. Pastoor stated that the description of the X and Y axes in the diagram (see Figure 2 - exposure and toxicity, respectively) are independent of policy, but the placement of the dotted line is a policy decision. To aid in transparency, the toxicity value on the Y axis is the animal or human data, without the application of uncertainty factors. Dr. Pastoor noted that one must compare "like with like." That is, if human exposure in on the X axis, one needs human toxicity values on the Y axis. Considerations of the type of data used on the X and Y axis also affect the placement of the dotted line defining exposures of concern. In the pyrethroid example, the toxicity data were from animal and *in vitro* studies so the diagonal dotted line was placed at an exposure: toxicity ratio of 1:100. Consideration of uncertainty factors or margin of exposure can be considered in the data plotted on the Y axis or in the placement of the diagonal dotted line.

The HESI RISK21 Quantitative Key Events Dose Response Framework (Q-KEDRF), Dr. Ted Simon

ABSTRACT

Advancing the existing MOA / Human Relevance Framework (HRF) and Key Events / Dose-Response Framework (KEDRF) (see Figure 3) to make best use of quantitative dose-response and timing information for Key Events produced the Quantitative Key Events / Dose-Response Framework (Q-KEDRF). The Q-KEDRF provides a structured quantitative approach for systematic examination of the dose-response and timing of Key Events from the initial dose of a bioactive agent to the potential adverse outcome. Two concepts are noted as aids to increasing the understanding of MOA - Associative Events and Modulating Factors. These concepts are illustrated using two case studies; 1) cholinesterase inhibition by the pesticide chlorpyrifos, which illustrates the necessity of considering quantitative dose-response information when assessing the effect of a Modulating Factor - here, enzyme polymorphisms in humans, and 2) estrogen-induced uterotrophic response in rodents, which demonstrates how quantitative dose-response modeling for Key Events, the understanding of temporal relationships between Key

Events, and a counterfactual examination of hypothesized Key Events can determine whether they are Associative Events or true Key Events.

DISCUSSION:

A panelist commented that in thinking about thresholds, it is best to focus on biology over statistics, thinking first about what biology could give rise to a threshold. Dr. Simon agreed, and added that one should think in terms of the steepness of the dose-response curve rather than threshold/no-threshold. Another panel member suggested that one could frame the question of quantitative relevance to humans given the MOA, in terms of what data are needed to characterize the interspecies differences, such as physiologically-based pharmacokinetic (PBPK) models regarding the active metabolite. Dr. Simon agreed, noting that the tables shown in his presentation are a work in progress.

Another panelist asked about the modulating factors and how could they be included into the framework. Dr. Simon responded that his example did include this information. For example, obesity was noted as a possible modulating factor for cancer susceptibility. Dr. Simon suggested that fully including the modulating factors in the analysis would be a substantial effort, but agreed that such factors represent an important consideration. The panel member agreed that if obesity were a factor, it would be a huge effort to address it quantitatively, but that it could be highly relevant. Another panel member suggested that additional discussion of approaches for including modulating factors is needed.

Case Study Discussions

Three new case studies (one with two parts) were presented. Panel input was sought on the utility of the methods to address specific problem formulations, and on areas for additional development. Inclusion of a method or case study in the framework as an illustration of a useful technique does not imply panel acceptance of the chemical-specific outcome. All case study presentations are available at http://www.allianceforrisk.org/Workshop/WS6/WS6casestudies.html.

Table 1. Workshop VI-Summary of Case Study Discussions

New Case Studies	
Case Study: Endogenous Chemical Risk Assessment: Formaldehyde as a Case Example	**Authored by: Robinan Gentry; Tom Starr; Jim Swenberg**
The purpose of this case study is to discuss endogenous and exogenous formaldehyde DNA adducts and the quantitative implications of adduct formation using two approaches that accommodate endogenous production of formaldehyde: 1) a “bottom up” approach; and 2) a biologically-based dose-response (BBDR) model that includes computational fluid dynamic (CFD) modeling to describe the flux of formaldehyde into tissues. The BBDR model includes parameters linked to two MOAs proposed for tumor development. The “bottom up” approach	

uses a simple linear model (Starr and Swenberg 2013) that can be extrapolated upward from background (endogenous) exposure and response levels, rather than downward from the observable response range, and was described by the authors as a bounding approach or reality check based on conservative assumptions.

Due to the complexity of this case study, the discussion was structured as three major presentations on different aspects of the assessment, with each presentation followed by clarifying questions from the panel, and then a final panel discussion on the overall approach.

(1) Dr. James Swenberg made the first presentation, entitled "Endogenous Chemical Risk Assessment: Formaldehyde as a Case Example."

A panel member recommended that the term "background" be avoided, and instead the terms endogenous versus exogenous exposure should be used. In response to a question from an audience member about the possibility that the adducts identified as being endogenous might instead reflect a technical artifact, Dr. Swenberg stated that there is no question that the endogenous adducts were present in vivo; the same method is used for simultaneously measuring endogenous and exogenous adducts in a DNA sample. Dr. Swenberg further noted an issue raised previously about using a DNA matrix versus using analytical standards to create standard curves. In the case of a DNA matrix, in the development of standard curves for analytical quantitation using DNA and [$^{13}CD_2$]-hydroxymethyl dG as the analyte, no difference was found when compared to analytical standards. Another audience member expressed concern about adducts resulting from processes such as peroxidation or reaction with buffer. Dr. Swenberg noted that his laboratory has rigorously examined this issue and found that no problems were present in their research. The results of this research are being prepared in a manuscript for publication. The contributions from buffers were less than 1% of the ever present endogenous amounts. Dr. Swenberg further replied that these processes (e.g., peroxidation) also occur in the cell, and for one of the adducts (formyl lysine), there are consistent measurements of exogenous and endogenous N6-formyllysine and N^2hydroxymethyl dG between two different laboratories and two different adducts; the results are not just from one research group.

Dr. Swenberg noted that one of the reasons that adducts are referred to as biomarkers of exposure, with mutations being a biomarker of effect, is that adducts vary widely in the degree to which they accumulate, result in mutations, and thus have different implications for cancer. A panel member suggested that it may not be useful to differentiate whether an adduct is a biomarker of exposure versus effect. The NAS has stated that an adduct may be a biomarker of either exposure or effect; it depends on the MOA (NAS, 1991a, 1991b).

In response to a panelist question, Dr. Swenberg noted that they are currently evaluating which data set(s) would be appropriate as the basis for low dose dose-response assessment. A panelist noted that it is important to tie the analysis to what is causing the tumors and the level of increased risk. The panelist stated that exogenous exposure does not need to exceed endogenous levels in order to be meaningful; a 10% increase in endogenous levels could result in a meaningful increase in risk. Dr. Swenberg agreed, noting the need to bound the results for sensitive individuals, but not ultrasensitive (i.e., repair-deficient) people.

(2) Dr. Tom Starr presented the second talk, entitled "A Novel Bottom Up Approach to Bounding Potential Human Cancer Risks from Endogenous Chemicals." (See Figure 4)

In response to a panelist question, Dr. Starr noted that in the bottom up approach, all background risk is due to endogenous exposure, which is a worst-case assumption; if this is not true, the slope factor and the predicted risk from exogenous exposures would be lower. A panel member asked about the level of confidence related to the estimate of the steady state concentration of exogenous adducts (Cxss), noting that different values of Cxss were obtained from two different studies of different durations. The panel member also noted the importance of having a confidence interval for the estimate of the adduct half-life. Dr. Starr noted that the estimates could be affected by differences in study design, and that cytotoxicity at high dose levels and study duration could also impact the estimate. Dr. Starr also noted that they were working to develop confidence bounds on the estimated half-life. Dr. Starr agreed with a panel member suggestion that the bounding method he described, using adduct data as an example, could also be applied to baseline levels of cell division. It was also noted that the approach could apply to mutagenic events.

In response to an audience question about endogenous levels of adducts in bone marrow and some data indicating higher endogenous levels of adducts in this tissue, Dr. Swenberg noted that they are conducting additional studies in order to obtain better data on endogenous levels in controls. Dr. Starr noted that higher endogenous levels (higher C_0) would produce a lower slope factor in the bottom-up approach. Noting that endogenous levels of adducts were higher than exogenous levels, another audience member suggested that a nonlinear model would make more sense than a linear one. At the request of the chair, this audience member provided a written summary of her comments after the meeting. These comments are provided in Appendix 4.

A panel member expressed overall approval for the concept proposed by Dr. Starr, but noted two caveats for generalizing the approach. First, if more than one type of tumor is caused by the carcinogen, the analysis would need to calculate the risks separately for each tumor type and add them together. Second, if the tumor target in humans is not known, it may not be possible to determine which tumors to look at in the SEER cancer registry, since it is not possible to assume site concordance. Thus, if sufficient epidemiology data are not available, or tumor site concordance does not occur, the panel member suggested that this method should not be applied. Another panel member noted that MOA information, such as a biomarker correlated with the apical endpoint (tumor type), could be used if human epidemiology data are not available.

A panel member expressed approval of using the approach to evaluate comparative uncertainties, but stated that doing the bottom up approach in the absence of MOA information would have high uncertainty. Another panelist noted that sources of conservatism in the approach include the following: (1) attributing all of the disease incidence to endogenous exposure and (2) conservative estimate of endogenous exposure. This panelist then asked about sources of variability and non-conservatism. Dr. Starr noted that the uncertainty in the denominator of the slope that represents the concentration of endogenous adducts based on laboratory animal data, is higher than the uncertainty in the numerator that represents the probability of the cancers of

interest in the US population. This is because the former is based on a few hundred animals, compared to the far larger number of specific cancer deaths in the US population. Population variability was expressed in terms of the standard error of the mean, rather than attempting to characterize sensitive subgroups. A panel member suggested that if there is some factor, such as a latent variable that affects the propensity to get cancer from the endogenous exposure, this could contribute a non-conservative aspect to the analysis. Although such factors are not addressed in "standard" top-down risk assessment, a panel member suggested that one could account for such a variable if it is known. Another panel member suggested that one could use the "population attributable risk fraction" or a probability density function to evaluate what portion of the background rate is attributable to endogenous exposure.

(3) *Dr. Jeff Schroeter presented the third talk, entitled Using CFD Modeling and Dosimetry as a Framework to Incorporate Endogenous Formation into a Chemical Assessment."* Dr. Schroeter presented work on constructing a biologically-based dose response (BBDR) model comprised of a computational fluid dynamic (CFD) model to characterize airflow and formaldehyde uptake in anatomically accurate renderings of the upper respiratory tract in rats and humans coupled with a physiologically-based pharmacokinetic (PBPK) model of tissue-phase reactions including endogenous production of formaldehyde. A key finding from BBDR model was that, when exogenous exposure is in the low ppb range, endogenous levels create a back-pressure, decreasing uptake into the tissue.

In response to panel questions, it was noted that the decreased absorption is the net effect, reflecting the sum of absorption and desorption, and so the number of cells at risk would decrease. However, different areas of the nose would have different levels of exposure, due to localized airflow profiles, and this specification can be quantified using the BBDR model.

General Discussion:

After all of the author presentations, the panel discussed the overall case study and generalizable lessons from the effort. Panel members suggested that the authors lay out the range of different approaches for the formaldehyde risk assessment, comparing the uncertainties associated with each method, to aid in decision making. An author replied that this is part of the plan, although refinement of the target tissue dosimetry is needed prior to revising the BBDR. It was noted that BBDR models increase knowledge about the sources of uncertainty, but that uncertainties need to be considered in a comparative context. Another panel member suggested following a similar approach with other chemicals for which the parent or metabolite is formed endogenously, such as vinyl chloride, acetaldehyde, and ethylene oxide. Although the toxicokinetics differ among these chemicals, a generalizable method needs to be capable of being applied to chemicals with toxicokinetic differences. Panel members suggested that the authors identify the steps in their approach and how they could be applied to other steps, as well as the generalizable learnings.

In summary, the panel supported carrying the method forward. The panel noted that the bottom-up approach appears to be conservative and is conceptually straightforward. It is biologically based and uses human data in part. Data on cell replication rates or mutations could be used

instead of adduct data, depending on the MOA and data availability. The authors did a nice job of laying out the uncertainties in a comparative context in work to date. Caveats and areas for future development include the potential need to consider other tumors, and if relying on animal data for identifying the targets, it is important to consider site concordance, as well as connecting target tissue dose with the apical effect. It may also be important to consider variability in endogenous levels in humans. Comparison of the predictions with epidemiological data could provide a useful reality check. Finally, it was suggested that the authors consider the implications of assuming a sublinear dose response in the endogenous exposure range, if this appears to be appropriate. A member of the audience suggested that the "bottom-up" approach could be used to evaluate chemicals for which the MOA includes genotoxicity at low doses and other mechanisms at higher doses, by thinking in terms of an endogenous mutation, as opposed to endogenous exposure.

Hypothesis-Based Weight of Evidence (Naphthalene as an Example)	**Authored by: Lorenz Rhomberg, Lisa Bailey**

The Hypothesis-Based Weight of Evidence (HBWoE) approach (Figure 5) has been applied to several chemicals, in addition to the naphthalene case study presented at the meeting. The approach is hypothesis-based in the sense that it emphasizes articulation of the proposed bases for the relevance of the data to the causal question at hand, specifying the logic and reasoning. The approach integrates all of the relevant data (epidemiology, animal toxicology, mechanistic, toxicokinetic, *etc.*), both positive and negative, in terms of quality and relevance to humans in a way that allows each data set to inform interpretation of the other. The approach further synthesizes all of the data to determine overall plausibility for causality in humans, considering uncertainties and inconsistencies in the data sets and ad hoc assumptions that may be required for some of the hypotheses put forth.

A panel member asked whether further modeling and dosimetry evaluation would help in evaluating the MOA. The panel member also noted that information is available on cells at risk for all of the species of interest. Dr. Rhomberg replied that PBPK models are available for only some of the species of interest – rats and humans, but not mice or monkeys.

In response to a panelist question as to whether the method explains why mice get lung tumors after exposure to naphthalene, and whether the tumors are relevant to humans, Dr. Rhomberg stated that that the issue of mouse lung tumors would need to be addressed separately; the current work focused on rat nasal tumors. The panelist continued by noting that the nose contains every cell type that the lower respiratory tract contains, and suggested that a computational fluid dynamic (CFD) model may be needed to explain the differences between rats and mice. Although such a model has not been implemented yet for naphthalene, data are available on the mouse mesh and airflow, and information on metabolic capacity can be added.

A panel member commended the authors on the broader aspect of the case study illustrating the use of WOE analysis to design a research strategy, but asked why the case study had so much emphasis on explaining why the cytotoxicity in the nose of mice did not lead to tumors. Dr. Rhomberg replied that part of the reason that the research team pursued the issue of the mouse

nose was because they expected questions about why the cytotoxicity did not result in tumor formation. He also noted that this evaluation was a way to test the hypothesis regarding a key role of cytotoxicity. Cytotoxicity is clearly not sufficient for tumor development, but that statement just restates observation, without explaining it. A key component of the method is working through explanations of both why things happened and also why they did not. The panelist replied that this does not necessarily argue against cytotoxicity being a key event; early key events in a MOA can occur without late key events developing. The panelist noted that it would be important from the risk assessment perspective to determine whether there is some other key event that is relevant to humans; it is important to characterize the limiting aspects of the dose-response (i.e., these are the key events) and what is needed to scale to humans. Two additional panelists suggested quantitative explanations for the difference between rats and mice. One suggested that a labeling index study be conducted to determine whether the species differ quantitatively in the number of additional cell divisions. Another panelist suggested that the differences related to differences in internal dose.

In further discussion in response to an audience member question about evaluating MOA, panel members noted that box checking approaches should be avoided; instead, one needs to understand the key events of the pathogenesis process. They also noted the importance of bringing the research and risk assessment communities together to understand MOAs /adverse outcome pathways, and in defining the minimal dataset to determine an MOA. It is anticipated that the AOP/MOA wiki that Dr. Meek mentioned in her earlier presentation will contribute to increasing interface and common understanding. It is also anticipated that progress in the Human Toxome project described by Dr. Hartung will help with these efforts. It is important to revisit criteria as the understanding of the relevant biology evolves.

Noting that an important part of the process is identifying ad hoc assumptions, a panel member asked for additional guidance on how to identify such assumptions. Dr. Rhomberg agreed that ad hoc assumptions can be hard to recognize, and stated that an ad hoc assumption is one that is invoked to explain an observed phenomenon, and would not be included in the hypothesis except for the need to explain certain data. The assumption may be correct, but the need to invoke it weakens the hypothesis.

With regard to generalizability, Dr. Rhomberg noted that the HBWoE approach has been applied for five chemicals, and has been used to address a number of complicated questions. For example, it has been used to consider such issues as evaluation of the formaldehyde epidemiological data and inconsistent results regarding the relationship between formaldehyde exposure and leukemia, and evaluation of whether chlorpyrifos exposure can cause neurodevelopmental effects at doses below those causing cholinesterase inhibition. Panel members recommended that it would be useful in this workshop series to look across chemicals at multiple examples to further evaluate the method.

In response to a question from an audience member about methods for comparing alternative "accounts," Dr. Rhomberg stated that they have avoided scoring systems, since once a system is codified, its scope is narrowed. Rules can go out of date and give a false sense of precision.

The panel discussed the similarities and differences between the HBWoE approach and the IPCS/ILSI approach for evaluating MOA and weight of evidence. It was noted that the two approaches are very similar, but HBWoE focuses more on what is needed to accommodate different explanations. A panel member observed that the MOA approach has evolved, and most of the stated differences between the two approaches noted in the authors' presentation no longer apply. The biggest difference is that the IPCS/ILSI MOA approach has been simplified to facilitate use, and the HBWoE approach now needs to be simplified for easier understanding and use. A panel member recommended that presentations of the outcome of HBWoE analysis should list what is consistent with the hypothesis, what is inconsistent, and plausibly describe the issues. Dose-response considerations should be integrated at an early stage. For both the HBWoE and the IPCS/ILSI approaches, key aspects include being transparent about the issues, and showing the relative weight of the alternative hypotheses/accounts. Both approaches differ from systematic review methods in that they focus more on how to integrate data after critical information has been identified. The HBWoE approach also focuses on explaining the observed results, noting the competing arguments and being clear about the implications of those arguments. A panel member noted the importance of applying such approaches in case studies, as a basis to evolve them based on those learnings, and codifying the learnings. The authors noted that their group is the only one that has applied the HBWoE system, although Borgert et al. (2011) did something similar. In order to apply this approach, expertise is needed to interpret the different types of input data, with enough understanding of the biology to consider the consequences of an "account." For example, it is important to think about how an "account" relates not only to the target, but to other endpoints. An important aspect is knowing how to find negative data. Panel members recommended that this work be integrated with the AOP/MOA wiki, and that aspects of the HBWoE approach be be included in the templates for the AOP/MOA wiki.

Panel members noted the importance of time course data, both for characterizing kinetics, and for considering the time required for lesions to develop. For example, one needs to take into account the time for remodeling of the perturbed epithelium of the respiratory tract.

For further applications of the HBWoE system, some panel members suggested using REACH or TSCA data sets to save time on the initial compilation of the data, although another panelist noted that the typical REACH datasets do not address MOA. In response to panelist questions, Dr. Rhomberg noted that using formal quantitative methods to analyze the case study data sets would be challenging, due to their large size, but that the diagram for a Bayesian analysis would be feasible to develop based on the existing work.

Building on Dr. Pastoor's presentation on Risk21, which integrates exposure and toxicity information to characterize an exposure as being of concern, of no concern, or somewhere in between, the panel and Dr. Rhomberg discussed how the Risk21 approach relates to the MOA and HBWoE approaches. Dr. Rhomberg stated that he sees HBWoE as being a higher-tier approach in the Risk21 scheme. Alternative interpretations could result in different probability density functions on the Risk21 map, and assessors could then evaluate whether it is worth investing resources to resolve the difference. A panel member noted that the initial work on MOA was for data-rich chemicals, but the results from data-rich chemicals are now being used

to inform the data-poor ones. Panel members also noted the move towards problem formulation, and considering relevant exposure levels when designing toxicity studies; the need for new test guidelines was noted. Panel member noted that testing strategies were originally designed to go to high doses in order to maximize sensitivity, but a panel member noted that the issue with such extrapolation is that biology is not linear and high doses cannot be used to make up for low statistical power. Instead, one panel member suggested that MOA analyses and genomics can be used to make up for lower statistical power, although another expressed a desire for better validation of the utility of toxicogenomics. A third panelist suggested that we need to change the definition of risk to something more analogous to the clinical setting, where predictions can be made as to whether a disease will advance based on the observed perturbation related to the disease. Rusty Thomas has found impressive correlations between dose-response data for genomic changes in *in vivo* studies and apical endpoints. Relevant lessons learned from Health Canada's project prioritizing 23,000 chemicals on its Domestic Substances List (DSL) include the use of tiered approaches, the importance of a focus on exposure, and the need for methods applicable to data-poor chemicals, such as the use of surrogates.

Interpretation of 24-hour sampling data. Case Study A. Texas Commission on Environmental Quality Approach	**Authored by: Roberta L. Grant; Allison Jenkins; Joseph (Kip) Haney**

The TCEQ develops Air Monitoring Comparison Values (AMCVs) to evaluate air monitoring data for potential health effects for specific averaging times: 1-hr and 24-hr measured air concentrations and calculated annual averages. The TCEQ asked the panel to review the methodology for developing 24-hr AMCVs. The basic analytical steps for developing 24-hr AMCVs are similar to procedures for developing 1-hr and annual (chronic) AMCVs for chemicals with a threshold dose-response based on published guidelines (TCEQ 2012). Since the 24-hr AMCV is specific to the exposure period and health effect(s) being considered, it may be used to conduct a risk assessment in combination with 1-hr and annual AMCVs, although concerns about intermittent exposure should be addressed by evaluating air monitoring data.

Panel members suggested additional sources for acute exposure limits that TCEQ could consider using, including the following: Acute Exposure Guidance Limits (AEGLs; longest duration is 8 hours), and Provisional Advisory Levels (PALs) developed by EPA's National Homeland Security Research Center (NSRC), which include 24-hour values.

In response to a panelist question as to whether values are needed for other short-term durations, Dr. Grant replied that the TCEQ currently develops 1-hour and chronic values based on the needs of its air permit review program and plans to develop 24-hr values for review of 24-hr air monitoring samples. The need for 24-h AMCVs is based on the fact that these values will allow a more full evaluation of air monitoring data and that the general public frequently compares 24-hr air monitoring data to chronic values. Although exceedances of the chronic limit are rare, it is useful to have 24-h AMCVs in cases where 24-hr measured concentrations are above chronic values, to address public concerns. The 24-hr AMCV was developed to aid in evaluating the impact of spikes of higher concentration above the chronic values. Such spikes are usually due to other-than-normal operations, such as a leaking valve.

A panel member expressed concern with using Haber's conjecture for duration adjustments, especially for irritants. The panelist noted that Haber's conjecture, including the ten Berge modification, is based on high concentration, short duration data for acute lethality. The approach assumes static toxicokinetics and toxicodynamics. Dr. Grant replied that issues related to the proper application of Haber's conjecture and when other approaches should be used are addressed in detail in the TCEQ guidelines and in the supporting chemical-specific Development Support Document (DSD).

The panelist then asked whether kinetic models are used when available. The TCEQ uses such model results when available, verified, applicable to the toxicity assessment, consistent with time and resource constraints, and uses all chemical-specific data in making an assessment. The panel member suggested that available PBPK models could have been used for the three example chemicals as they each have published model structures. The case study authors replied that these models had not been used because they had not been validated or did not predict the endpoint of concern. The panel member recommended that using the amount of parent chemical in the blood provides a better basis for the assessment than using external exposure, and that this can be done without the more challenging development of a full understanding of which metabolite causes an effect and description of metabolite levels.

In further discussion, Dr. Grant noted that TCEQ does not have the resources to do its own modeling. Panel members noted that it would be useful to have people who could advise agencies on modeling, and that ATSDR has a published papers on developing a PBPK toolkit that provides general model structures and vetted code, and these will become available on-line shortly (Mumtaz et al., 2012; Ruiz et al. 2011; Ruiz et al., 2010).

An audience member (Dr. Rhomberg) noted that he had published a paper evaluating whether Haber's conjecture and the ten Berge modification can be derived based on fundamental toxicokinetics and simple toxicodynamic assumptions (Rhomberg, 2009). The paper found that both equations are special cases of a larger relationship that applies to all chemicals at different durations, with the elimination half-life being a key determinant. Dr. Rhomberg suggested that the TCEQ approach method in general is appropriate in considering persistence of the chemical and of the effect, but suggested that the Rhomberg framework be used to generalize and rationalize the method. A panel member noted that Fred Miller also has a useful paper on this topic (Miller et al., 2000), and John Doull had a special issue on the topic of Haber's rule (Rozman and Doull, 2001a, 2001b; Doull and Rozman, 2000; Rozman and Doull, 2000). The panel member also stated that physiologically-based methods are better than empirical ones, noting that computational fluid dynamic modeling captures relevant information on airflow and anatomy.

Another panelist compared the TCEQ values and ATSDR values and found that the TCEQ values are typically higher. The authors replied that differences in acute values relate to factors such as differences in duration. ATSDR values are designed for durations up to 14 days whereas the acute AMCVs are based on a single 1-hr or 24-hr exposure, respectively. Scientific judgment on uncertainty factors also may be different. A panelist recommended that information on

comparison with existing values should be reflected in the documentation. The assumptions regarding uncertainty factors and duration context can be backed out, providing additional perspective on the 24-hour AMCV. Dr. Grant noted the chemical-specific DSDs often include a comparison with other agency values and a discussion on why the values may be different.

In response to a panel member question about the strategy if a 24-hour AMCV were exceeded, Dr. Grant stated the regional TCEQ office would be contacted to find out the reason. The exceedance(s) would not be interpreted as meaning that health effects would occur, since the 24-hour AMCVs are below levels where health effects would be expected. TCEQ would evaluate the exposures in the context of considering the 24-hour air monitoring data, and how the 24-hour AMCV was derived. The TCEQ would work with the regional office and with the relevant source(s) to address the issue. If the frequency of exceedances of the 24-hr AMCV and the peaks were high enough, the annual average limit would be exceeded. The TCEQ evaluates ambient air data based on rolling annual averages to identify upward trends relative to the annual average AMCV.

Interpretation of 24-hour sampling data. Case Study B. Ontario Ministry of the Environment (MOE) Approach	**Authored by: Denis Jugloff; Julie Schroeder**

The MOE develops science-based ambient air quality criteria, or AAQCs, at concentrations that are protective against effects that may occur during continuous lifetime exposures to a contaminant. A 24 hour averaging time is used for to allow evaluation of 24-hour sampling data, and may be assigned to an AAQC based on one of two methods (see Figure 6). Method # 1 is to set a 24-hour averaging time for an effect that may be caused by short-term exposures (e.g., reproduction, development), or if daily variation of the mean exposure, afforded by an annual average, is not considered acceptable. Method #2 is to first set an annual AAQC for contaminants causing adverse effects after chronic exposures, and then use a meteorological conversion factor to convert the annual AAQC to a 24-hour AAQC. For both methods, if the 24-hr AAQC is met, then no adverse effects are expected over continuous lifetime exposure, and it is expected to be protective of other effects, including high-dose effects over short-term acute exposures. Thus, the advantage of using a 24-hour averaging time is that it provides a single target that is protective of both short-term and long-term exposures. However, for Method #2, as the converted 24-hour AAQC is not directly linked to an effect and instead provides an indication whether the effects-based annual AAQC may be exceeded, the converted 24-hour AAQC is not suitable for use in risk assessments; rather, the original effects-based annual AAQC is to be used for risk assessments.

In response to panelist questions, Dr. Jugloff and his colleagues noted that although the MOE maintains an extensive ambient air monitoring network, data specific to a contaminant, source and/or location may be limited. Therefore, it may not always be possible to calculate a true annual average.

Regarding method #1, the panel questioned the need for extrapolating chronic AAQCs to acute exposures since the chronic values should already be protective of more sensitive acute effects, based on the way the value is derived. There was some concern from the panel as well regarding MOE's use of the qualitative acute toxicities in the quantitative development of several chronic

AAQCs, or regarding the issues of potential windows of susceptibility, such as the case of developmental effects from manganese exposure. A panel member explained this concern by saying that some effects can occur with short-term exposure, of course, and that regardless of whether this effect occurs over long or short time, a chronic number should be protective of short-term exposure, even if a different endpoint is used.

The authors replied that the manganese AAQC was derived quantitatively, but that the averaging was qualitative. In other words, the manganese AAQC was *quantitatively* derived from an endpoint of eye-hand coordination impairments in workers with several months to 15+ years exposure (average of 5.3 years), with applied uncertainty factors for inter-individual susceptibility (UF=10), database limitations specific to Mn speciation (UF=3), susceptibility of the developing nervous system (UF=3), and subchronic-to-chronic extrapolation (UF=3). However, there remained *qualitative* concerns that manganese is a known developmental neurotoxin, and that exposure to short-term peaks at key points in development might have life-long implications. Thus, for this latter qualitative issue, the authors had a remaining concern that if the manganese AAQC was assigned an annual average, short-term excursions could be an issue; thus, a 24-hour averaging time was assigned. The panel likened this concern to that expressed and resolved with methods for Occupational Exposure Limits (OELs). For OELs, exposure limits are based on 8-hr averages, but short term exposure limits or ceilings can be set if there is concern about short-term spikes.

Consistent with the intent of MOE, the panel thought that it is critical to look at what other effects happen at somewhat higher levels after acute exposure (and potentially described in the documentation of an AAQC), and that this needs to be clearly communicated. The panel also thought that when MOE contemplates different approaches that its rationale should be documented, especially since some approaches account for toxicokinetics and different lifestages. Overall, the panel felt that because uncertainty factors are used in the chronic AAQC that this accounts for some of MOE's issues associated with extrapolation to acute exposures. However, the panel noted that for certain chemicals, such as dioxin, a short term exposure could ultimately be the basis for a long term AAQC due to the long half-life, and for this reason that it is important to include toxicokinetics.

Panel members expressed further concern about using the first approach (based on short term effects) for persistent bioaccumulative chemicals, even if the endpoint was developmental toxicity. A panel member also noted that response to respiratory irritation takes some time to evolve (i.e., time for lesions to develop in remodeling response to epithelial perturbation), which can make it problematic to use the acute data. It is also important to recognize that there are differences between acute and chronic irritation, and effects may occur from chronic exposure that would not occur following acute exposure. For example, acute exposure limits would not be protective for chronic exposure; repeated exposure to levels that cause mild effects can cause cumulative damage that results in severe effects after a longer exposure duration. This panel member suggested that an adjustment factor could be used to account for limitations in the data. In considering the approach and the potential need for an adjustment factor, one needs to consider whether the acute and chronic values are based on the same endpoint. If they are based on different endpoints, the potential need for an adjustment factor should be considered.

The authors restated the panel discussion as recommending that if there are either no data for a secondary short-term effect such as developmental toxicity, or if the data indicate that effect levels are close to or just above the point of departure for the critical effect, then MOE needs to consider whether short-term spikes will evoke the secondary short-term effect. If one doesn't have the data, the issue of the potential impact of a short-term spike can be considered based on whether an uncertainty factor for database deficiency (UF_D) was used in developing the chronic AAQC. MOE would then determine whether different endpoints exist between acute and chronic exposures, and whether developmental effects need to be considered separately or whether a separate short-term value should be developed.

Another panel member mentioned preference for a short term value based on a health effect appropriate to that averaging time. For 24 hour averaging time the endpoint should be based on acute effects and that this requires different levels based on different durations.

A member of the audience noted that for irritant effects, the half-life of repair is slow relative to the elimination half-life of the chemical; the issue is which of these half-lives drives the effect. The audience member recommended that questions about averaging time could be better handled by looking at half-life of elimination for "discounting" previous exposures, rather than using rolling averages.

Another panel member had concern that two things could counter-weight the approach: 1. Is there a sufficient database on the potential for developmental toxicity; this can be considered based on whether the chronic value included a database uncertainty factor. 2. There is the potential that a single or a few short-term exposures could results in the annual average being exceeded. The authors clarified that when there are no data available, UFs are still appropriate. The discussion relates to if there *is* a short-term concern, then one needs a 24-hr AAQC for this other short-term effect.

Another panel member mentioned that a short-term window of exposure is not just an issue of whether a chronic value takes into account the potential for reproductive/developmental effects. One needs to look at the overall data and whether they expect to see the effects. If the concern is a window of susceptibility, then it needs bounding. If one has the reproductive/developmental data, how protective will this be for the window of exposure? These points need to be communicated. Another panel member thought that when there are concerns about short-term windows that these are appropriate to address, but maybe not in the averaging time, instead they should be addressed more in the development of a number specific to the duration.

Introducing the second approach, Dr. Jugloff noted that MOE typically does not have sufficient data to calculate a true annual average, and often has only one measurement. The essence of the approach is stating that if the single measurement is within a factor of five of the chronic value, one can make the judgment that the limit based on the annual average is met. This approach is assuming that the single measurement is representative of long-term exposures, rather than relating the 24-hour AAQC to short-term health effects. The factor of five is based on meteorology and topography, and although monitored data showed a range in the relationship

between averaging times depending on location, a ratio approximating 5:1 was observed in most cases, supporting the selection of a province-wide five-fold conversion factor.

A panel member recommended that this factor could be made more predictive by weighting it to where most of the Ontario population lives, and the relevant topology; although EPA uses the same factors as used by MOE in some situations, it also uses more specific adjustments for individual regional populations. Another panelist noted that the factor of five should be characterized as a high percentile of the distribution, not as an absolute maximum. The panelist also suggested that excursions could last more than 24 hours. Dr. Jugloff replied that there may be value in considering using 2-week values, analogous to ATSDR, but the 24-hour value was related to the sampling strategy. Panel members agreed that the second approach is a pragmatic one related to exposure assessment and exposure modeling, that aids in determining where additional resources should be invested.

A member of the audience recommended that the issue of episodes of higher exposures be considered using a more theoretical approach, recognizing that kinetic and dynamic changes related to the chemical exposure occur over time. From the exposure perspective, one can envision a moving average body burden that is weighted by half-life. From the dynamic perspective, one can imagine evaluating cumulative damage vs. repair. Based on these considerations, one can then evaluate when averaging time assumptions would be problematic. Panel members suggested case studies that have evaluated how internal dose varies with time, with the biological half-life being a key determinant. Specifically, Teeguarden's work on variability in internal dose of bisphenol A based on pharmacokinetic modeling (Teeguarden et al., 2011), and Price's work on chlorpyrifos (Hinderliter et al., 2011; Price et al., 2011) serve as illustrative examples. A panel member also noted the work of Saltzman (1996) showing that, with time-varying exposures, fluctuations lasting on the order of a ¼ half-life or less are averaged out and do not substantially affect the time course of the blood level. So, for chemicals with longer half-lives, finer resolution in the evaluation will not result in a different answer. Because the longest half-life is the key determinant, it is possible to obtain a useful approximation from simple models, without a full PBPK model. Panel members noted that this discussion highlights the utility of connecting exposure and pharmacokinetic modeling and that one could refine either aspect.

A panel member noted the utility of documenting and explaining the different limits and what they mean.

In summary, the panel had the following comments regarding the first approach (i.e., basing the 24-hour AAQC on effects from short-term exposures): Developmental toxicity is considered for chronic exposure limits. For developing 24-hour AAQCs, one needs to consider the chemical's toxicity profile and whether a health concern exists for short-term exceedances of a chronic limit based on the chemical's spectrum of effects, rather than applying uncertainty factors in a default manner. The panel also emphasized that it is important to look at what effects occur at exposure levels somewhat higher than the chronic limit. It is important to consider the toxicokinetics and toxicodynamics of the chemical of interest; some specific issues were noted in the above discussion related to the case study chemicals. Additional resources for addressing the issues

were also noted by the panel members and audience.

For the second approach (i.e., based on conversion from the annual AAQC), the panel agreed that the method provides a practical approach for developing 24-hour AAQCs. The panel recommended that MOE clarify that the issue is interpretation of a single measurement relative to the annual average, and that daily 24-hour averages are not available. The panel also recommended that MOE clarify the source of the factor of five used in the adjustment, and whether that factor is a percentile or maximum. Finally, it was noted that a region-specific model could be used, weighted towards where the provincial population is, rather than a data-based default factor of five.

References

ARA (Alliance for Risk Assessment). (2012). ARA Dose-Response Framework, Draft. Phase I: Problem Formulation and Scoping. Available online at: http://chemicalriskassessment.org/methods/.

Borgert CJ, Mihaich EM, Ortego LS, Bentley KS, Holmes CM, Levine SL, Becker RA. (2011). Hypothesis-driven weight of evidence framework for evaluating data within the US EPA's Endocrine Disruptor Screening Program. Regul Toxicol Pharmacol. Nov;61(2):185-91. doi: 10.1016/j.yrtph.2011.07.007.

Doull J, Rozman KK. (2000). Using Haber's law to define the margin of exposure.Toxicology. Aug 14;149(1):1-2

Hinderliter PM, Price PS, Bartels MJ, Timchalk C, Poet TS.(2011). Development of a source-to-outcome model for dietary exposures to insecticide residues: an example using chlorpyrifos. Regul Toxicol Pharmacol. Oct;61(1):82-92. doi: 10.1016/j.yrtph.2011.06.004.

IPCS (International Programme on Chemical Safety) (2007). IPCS framework for analysing the relevance of a cancer mode of action for humans and case studies. (http://www.who.int/ipcs/methods/harmonization/areas/cancer_mode.pdf)

Miller FJ, Schlosser PM, Janszen DB. (2000). Haber's rule: a special case in a family of curves relating concentration and duration of exposure to a fixed level of response for a given endpoint. Toxicology. 2000 Aug 14;149(1):21-34.

Mumtaz MM, Ray M, Crowell SR, Keys D, Fisher J, Ruiz P. (2012). Translational research to develop a human PBPK models tool kit-volatile organic compounds (VOCs). J Toxicol Environ Health A. 2012;75(1):6-24. doi: 10.1080/15287394.2012.625546.

National Research Council. (1991a). Human exposure assessment for airborne pollutants: Advances and opportunities. Washington DC. National Academy Press

National Research Council (1991b). Environmental epidemiology; v. 1. Public Health and hazardous wastes. Washington DC. National Academy Press.

NRC (National Research Council of the National Academy of Science). (2009). Science and decisions: Advancing risk assessment. National Research Council, National Academies Press, Washington, DC. AKA, “Silverbook”.

Price PS, Schnelle KD, Cleveland CB, Bartels MJ, Hinderliter PM, Timchalk C, Poet TS. (2011). Application of a source-to-outcome model for the assessment of health impacts from dietary exposures to insecticide residues. Regul Toxicol Pharmacol. 2011 Oct;61(1):23-31.

Rhomberg, Lorenz R. (2009). Uptake Kinetics, Species Differences, and the Determination of Equivalent Combinations of Air Concentration and Exposure Duration for Assessment of Acute Inhalation Toxicity, Human and Ecological Risk Assessment: An International Journal, 15: 6, 1099 — 1145: DOI: 10.1080/10807030903304724

Rozman KK, Doull J. (2001). Dose and time as variables of toxicity.Toxicology. 2000 Apr 3;144(1-3):169-78.

Rozman KK, Doull J. (2001a). Paracelsus, Haber and Arndt. Toxicology. 2001a Mar 7;160(1-3):191-6.

Rozman KK, Doull J. (2001b). The role of time as a quantifiable variable of toxicity and the experimental conditions when Haber's c x t product can be observed: implications for therapeutics. J Pharmacol Exp Ther. Mar;296(3):663-8.

Ruiz P, Fowler BA, Osterloh JD, Fisher J, Mumtaz M. (2010). Physiologically based pharmacokinetic (PBPK) tool kit for environmental pollutants--metals. SAR QSAR Environ Res. 2010 Oct;21(7-8):603-18. doi: 10.1080/1062936X.2010.528942.

Ruiz P, Ray M, Fisher J, Mumtaz M. (2011). Development of a Human Physiologically Based Pharmacokinetic (PBPK) Toolkit for Environmental Pollutants. Int J Mol Sci. 2011;12(11):7469-80. doi: 10.3390/ijms12117469. Epub 2011 Oct 31.

Saltzman BE. (1996). Assessment of health effects of fluctuating concentrations using simplified pharmacokinetic algorithms. J Air & Waste Manage Assoc 46:1022–34

Teeguarden JG, Calafat AM, Ye X, Doerge DR, Churchwell MI, Gunawan R, Graham MK. (2011). Twenty-four hour human urine and serum profiles of bisphenol a during high-dietary exposure. Toxicol Sci. Sep;123(1):48-57. doi: 10.1093/toxsci/kfr160.

Figures

Figure 1. Modified MOA Framework, from Presentation by Dr. Bette Meek

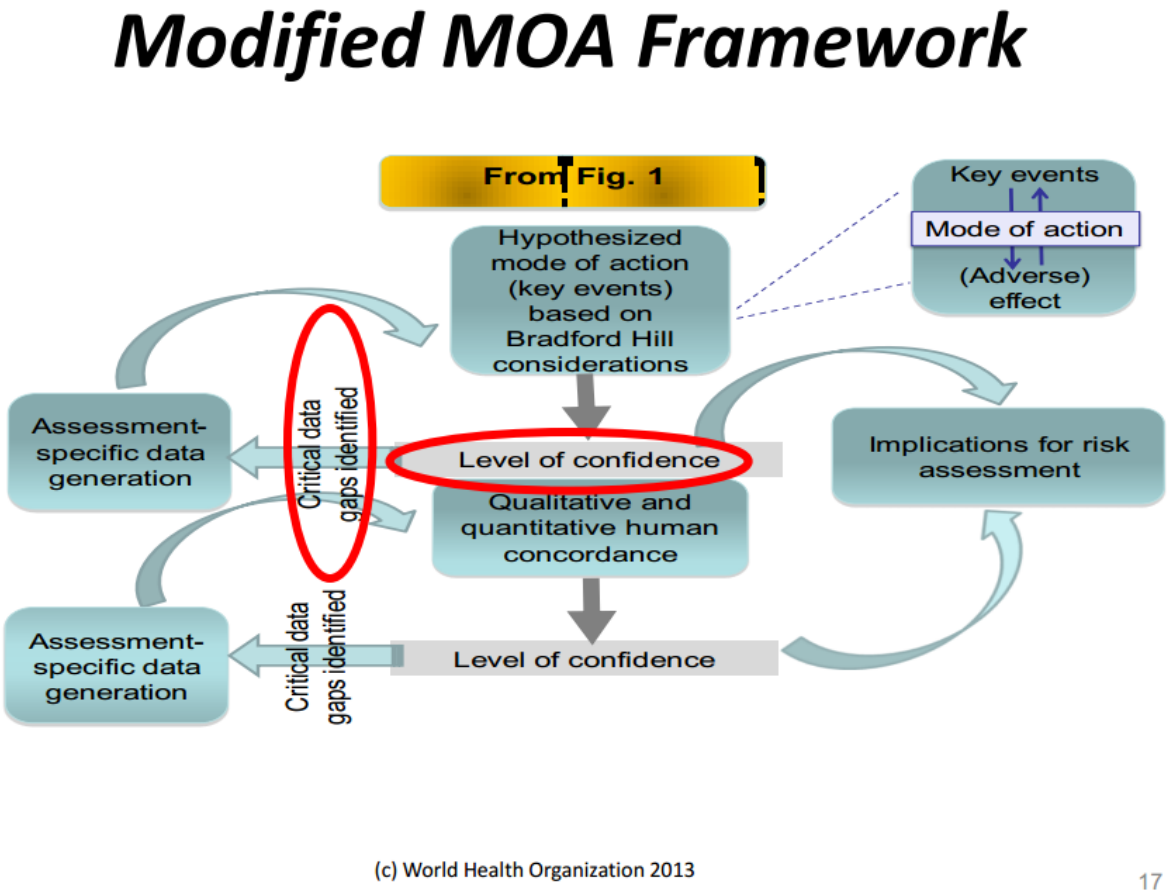


Figure 2. Sleeping Scenario: Tier 1, Illustrating Risk21 Approach for Integrating Exposure and Toxicity Information, from Presentation by Dr. Tim Pastoor

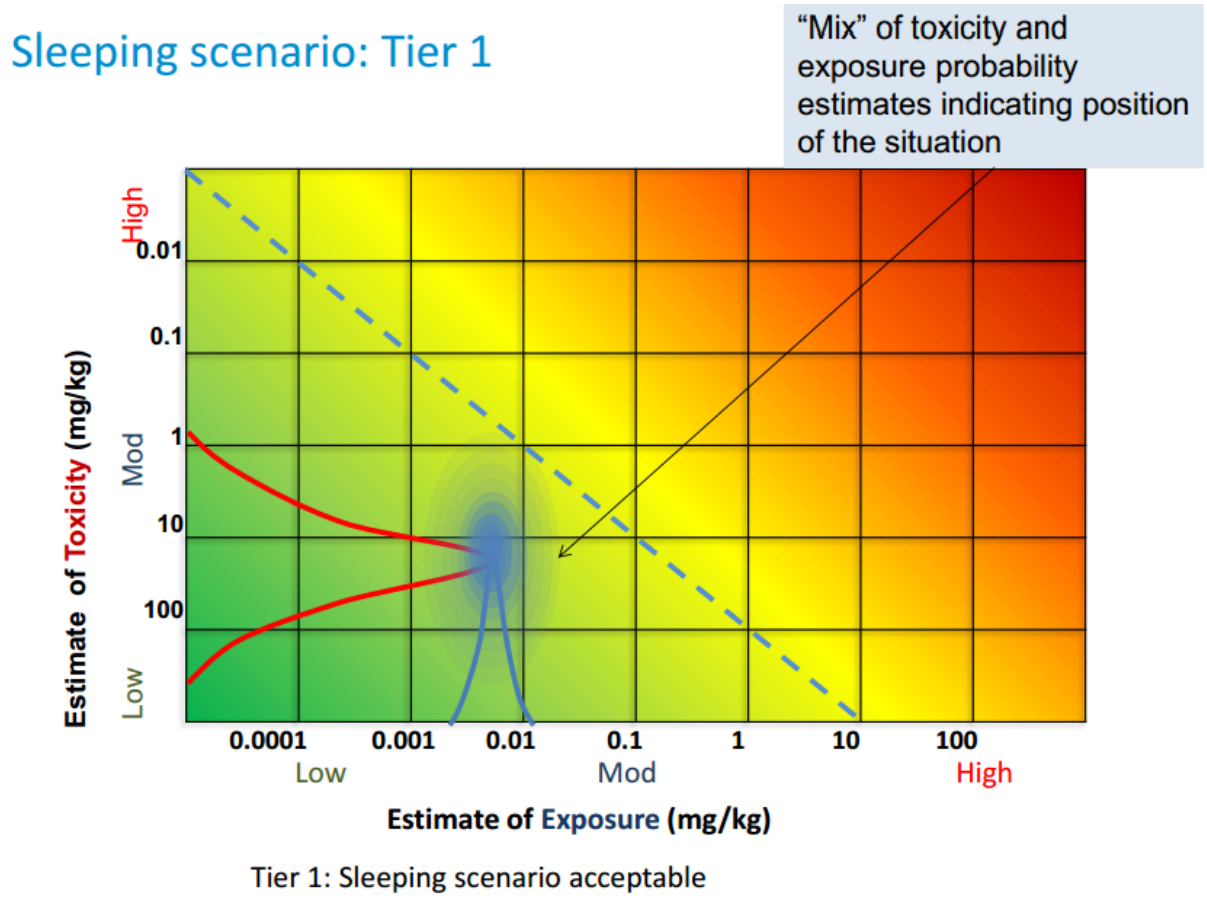


Figure 3. Quantitative Key Events/Dose-Response Framework (Q-KEDRF), from Presentation by Dr. Ted Simon

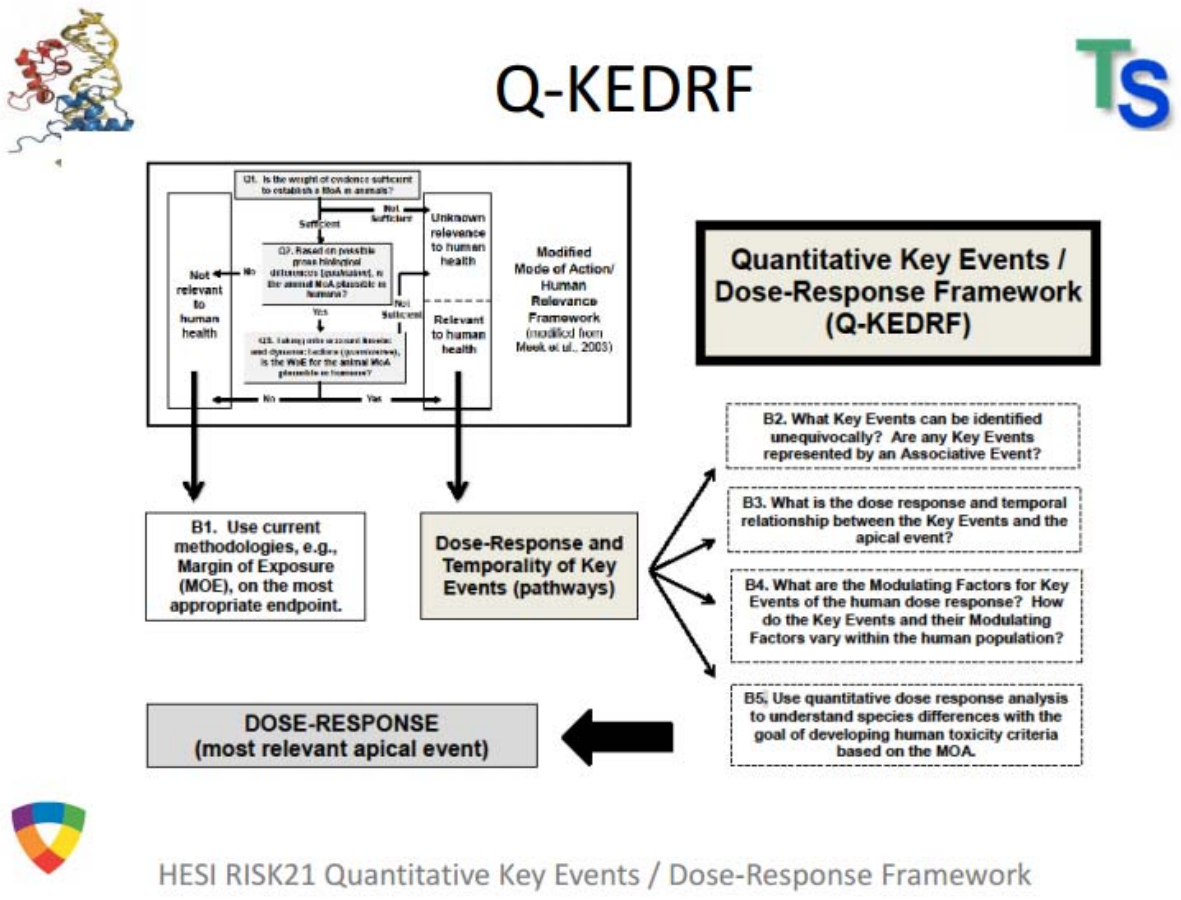


Figure 4. Bottom Up Approach Elements, from Presentation by Dr. Thomas Starr

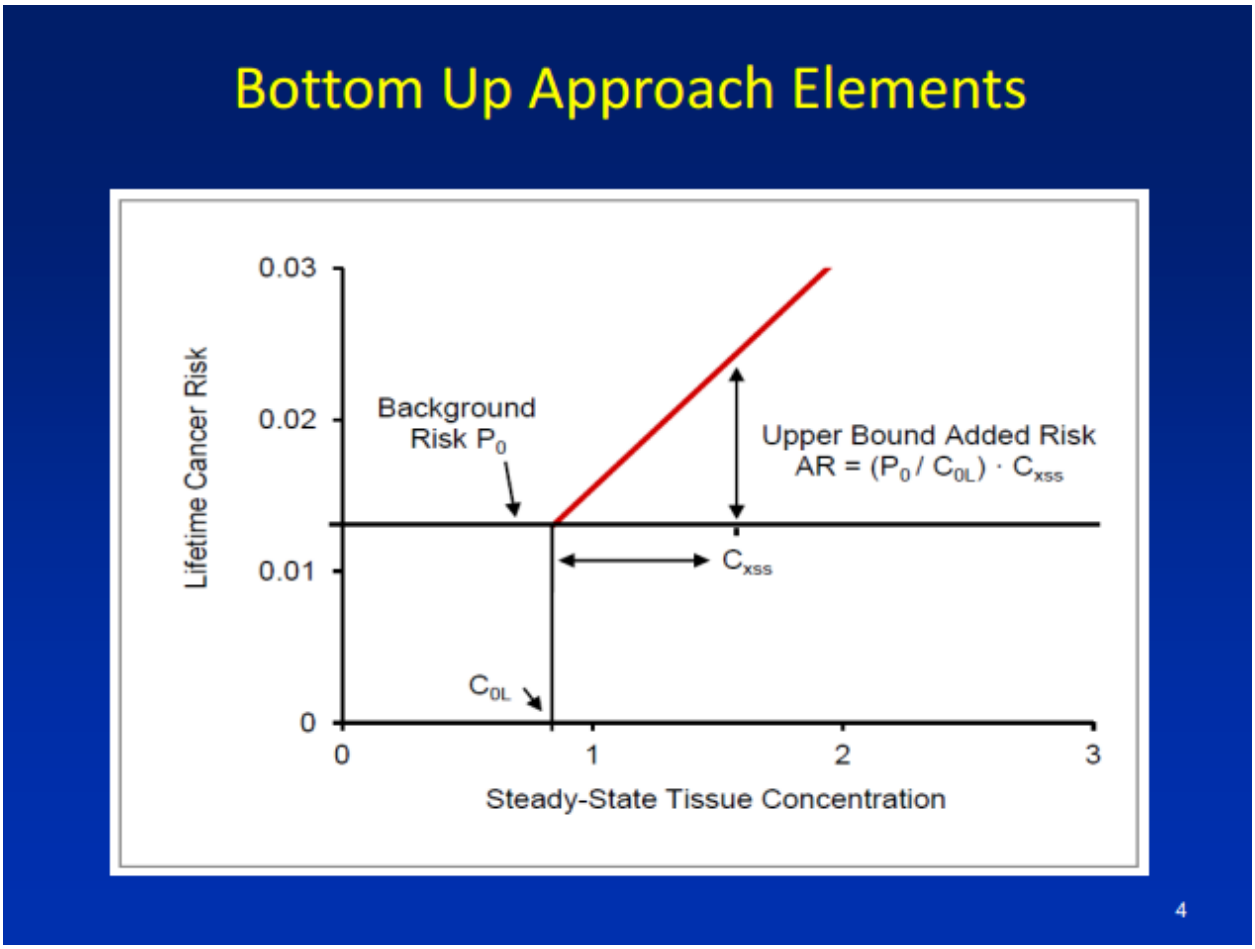


Figure 5. Hypothesis-Based Weight of Evidence (Napthalene as an Example), from Presentation by Dr. Lorenz Rhomberg

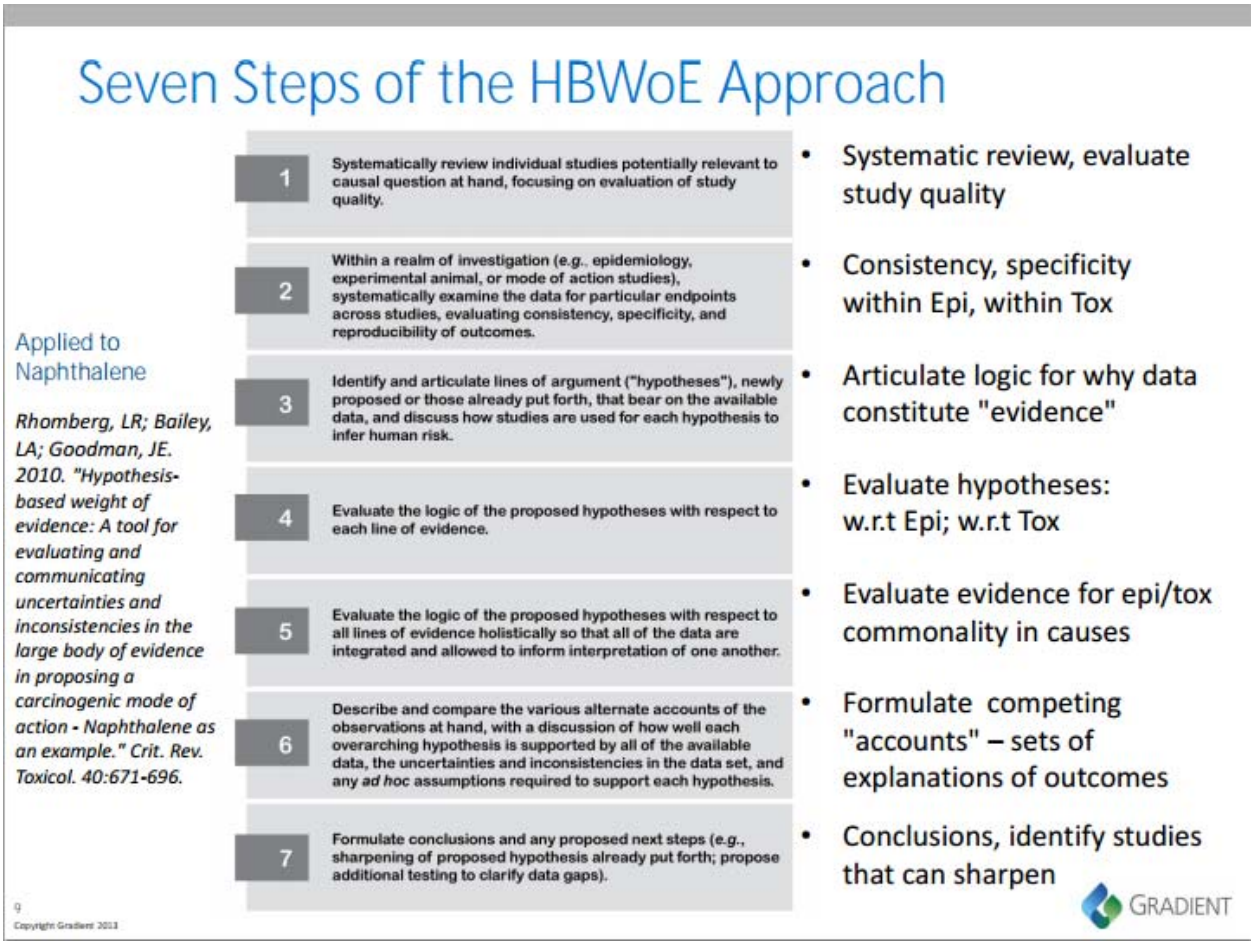
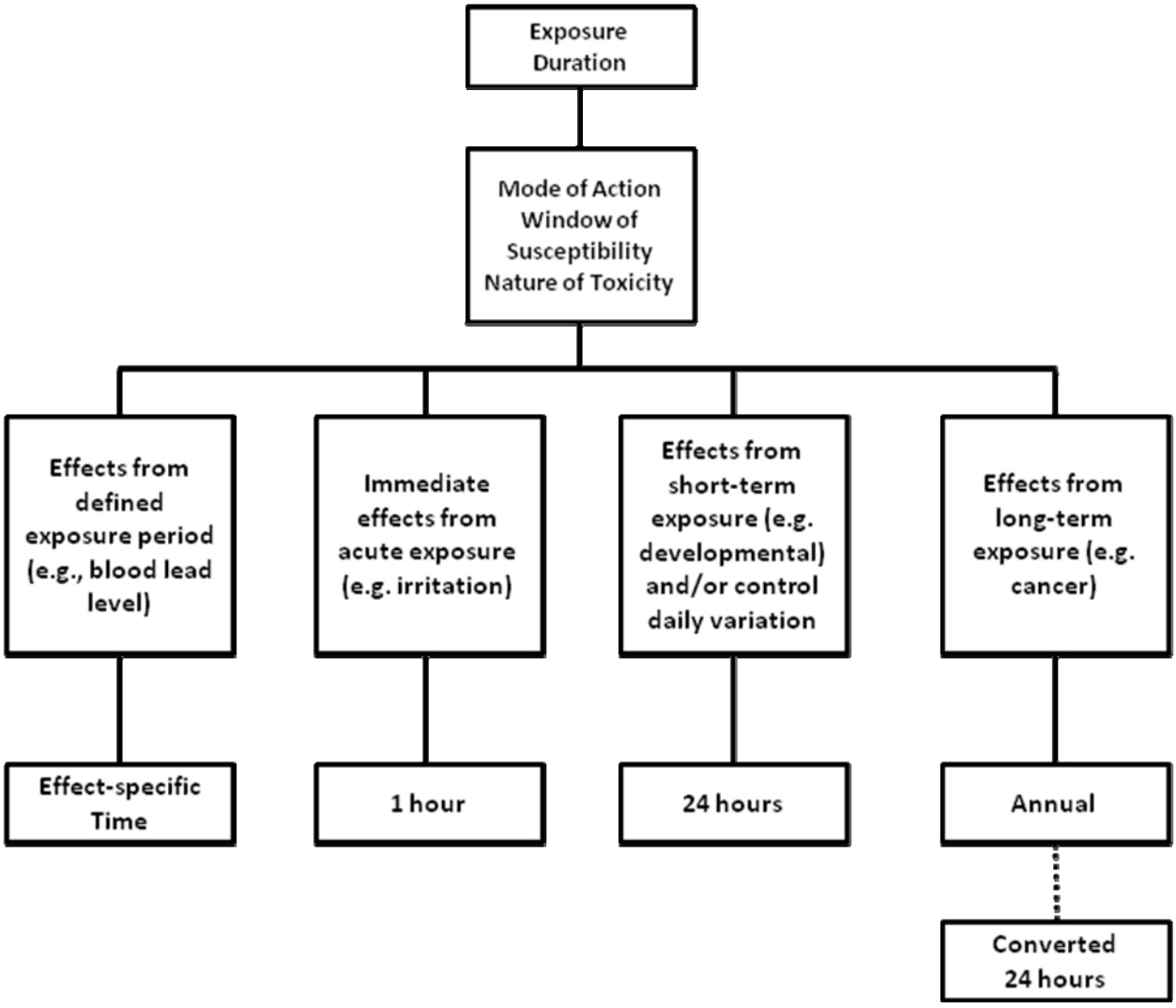


Figure 6. Ontario Ministry of the Environment Methods for Developing Ambient Air Quality Criteria (AAQCs)



Note: two different 24-hour AAQCs