Review Article

Journal of Applied Toxicology

Received: 2 November 2013, Revised: 28 November 2013, Accepted: 3 December 2013 Published online in Wiley Online Library: 10 February 2014

(wileyonlinelibrary.com) DOI 10.1002/jat.2984

Mode of action human relevance (species concordance) framework: Evolution of the Bradford Hill considerations and comparative analysis of weight of evidence

M. E. (Bette) Meek*, Christine M. Palermo, Ammie N. Bachman, Colin M. North and R. Jeffrey Lewis

ABSTRACT: The mode of action human relevance (MOA/HR) framework increases transparency in systematically considering data on MOA for end (adverse) effects and their relevance to humans. This framework continues to evolve as experience increases in its application. Though the MOA/HR framework is not designed to address the question of "how much information is enough" to support a hypothesized MOA in animals or its relevance to humans, its organizing construct has potential value in considering relative weight of evidence (WOE) among different cases and hypothesized MOA(s). This context is explored based on MOA analyses in published assessments to illustrate the relative extent of supporting data and their implications for dose–response analysis and involved comparisons for chemical assessments on trichloropropane, and carbon tetrachloride with several hypothesized MOA(s) for cancer. The WOE for each hypothesized MOA was summarized in narrative tables based on comparison and contrast of the extent and nature of the supporting database versus potentially inconsistent or missing information. The comparison was based on evolved Bradford Hill considerations rank ordered to reflect their relative contribution to WOE determinations of MOA taking into account increasing experience in their application internationally. This clarification of considerations for WOE determinations as a basis for comparative analysis is anticipated to contribute to increasing consistency in the application of MOA/HR analysis and potentially, transparency in separating science judgment from public policy considerations in regulatory risk assessment. Copyright © 2014. The Authors. Journal of Applied Toxicology Published by John Wiley & Sons Ltd.

Keywords: human relevance framework; mode of action; weight of evidence; key events; evolved Bradford Hill considerations

Introduction

The mode of action/human relevance (MOA/HR) framework is an analytical framework designed to increase transparency in the systematic consideration of the weight of evidence (WOE) of hypothesized MOA(s) for critical effects and their relevance to humans. It was developed in initiatives of the International Life Sciences Institute Risk Sciences Institute (ILSI RSI) and the International Programme on Chemical Safety (IPCS) and derives from earlier work on MOA by the US Environmental Protection Agency (USEPA) and IPCS (Sonich-Mullin *et al.*, 2001).

The development and evolution of the IPCS ILSI RSI MOA/HR framework, which has involved large numbers of scientists internationally, is described in several publications (Boobis *et al.*, 2006, 2008; Meek, 2008; Meek *et al.*, 2003; Seed *et al.*, 2005). Potential application in a broader range of relevant contexts has been considered more recently (Carmichael *et al.*, 2011; Meek and Klaunig, 2010). The framework has been illustrated by an increasing number of case studies ($n = 30$, currently), and is widely adopted in international and national guidance and assessments (Meek *et al.*, 2008), including those of the USEPA (Dellarco and Baetcke, 2005; Manibusan *et al.*, 2007; SAB, 1999, 2007; SAP, 2000; USEPA, 2005a). Building on this collective experience, the framework has been updated recently, to address uncertainty additionally and to extend its utility to emerging areas in toxicity testing and non-testing methods. The update includes incorporation within a roadmap, encouraging continuous refinement of fit-for-purpose testing strategies and risk assessment (Meek *et al.*, 2014).

In addition to increasing transparency through structured articulation of the evidence and uncertainties upon which conclusions are based, MOA/HR analysis also contributes to the transparent assimilation of all available data in both a risk assessment and research context. This is important because it facilitates identification of critical data needs and contributes to transparency in the separation of science judgment (i.e., weighting of options based on systematic consideration of available scientific support) from public health protection policy, the latter

*Correspondence to: M. E. (Bette) Meek, University of Ottawa, One Stewart Street, Suite 309, Ottawa, Ontario, Canada, K1N 6N5.
Email: bmeek@uottawa.ca

University of Ottawa, One Stewart Street, Suite 309, Ottawa, Ontario, K1N 6N5, Canada

sometimes involving embedded conservatism to increase public health protection.

Though the MOA/HR framework is not designed to address the question of "how much information is enough" to support a hypothesized MOA in animals or its relevance to humans, its organizing construct has value in considering relative WOE among different cases and hypothesized MOAs. Comparative WOE evaluation for MOA/HR analysis is illustrated as a basis to increase common understanding of the nature of transparency required to document the relative degree of confidence in supporting data for hypothesized MOAs. To demonstrate this approach, WOE for MOA/HR analysis in two published assessments (i.e., carbon tetrachloride and 1,2,3-trichloropropane [TCP]) (USEPA, 2009, 2010) is comparatively considered in the context of evolved Bradford Hill (B/H) considerations introduced here to promote better common understanding and consistency in use. The focus here is not on the conclusions of the assessments but rather, the utility of comparative analysis for WOE evaluation in MOA/HR analysis. These cases were specifically selected to exemplify varying degrees of WOE for several hypothesized MOA.

Methods And Results

Details of the updated MOA/HR framework are available elsewhere (Meek *et al*., 2014). Briefly, the WOE for a hypothesized MOA in animals is assessed based on considerations modified from those proposed by Bradford Hill (Hill, 1965) for assessment of causality in epidemiological studies. HR or species concordance is then systematically considered, taking into account more generic information such as anatomical, physiological and biochemical variations. If the WOE for the hypothesized MOA is sufficient and relevant to humans, implications for dose–response in humans are then considered in the context of kinetic and dynamic data. Delineation of the degree of confidence in the WOE for hypothesized MOAs is critical, as is the delination of critical research needs.

Establishing support for or rejection of a hypothesized MOA provides the foundation for subsequent considerations of dose–response, HR and estimates of risk. It involves (1) delineation of key events leading to the end (adverse) effect in a hypothesized MOA and (2) evaluation of all of the data to consider the extent of the supporting WOE for the hypothesized MOA. Importantly, if alternative MOA(s) are supported, these are evaluated with equal rigor in separate MOA/HR framework analyses. Ultimately, depending upon the application, there may be a need to draw a conclusion on the sufficiency of data supporting a MOA, to assess different risk management options. The comparative analysis of WOE was developed as a basis for increasing common understanding of the nature of transparency required to document the degree of confidence in the sufficiency of supporting data for hypothesized (potentially competing) MOAs.

A template for WOE analysis of MOA based on the evolved B/H considerations is presented in Table 1. In this approach, supporting data, inconsistent data and missing information are evaluated and tabulated in the context of the evolved B/H considerations presented here. The data in this table are considered in totality to assess the WOE for a MOA. In addition, the evidence can be used in a comparative manner to gain perspective on the relative degree of confidence that a hypothesized MOA is operative, based on the extent of supporting WOE compared to that for another postulated MOA for the same chemical or for the same MOA for other chemicals.

As illustrated in Table 1, WOE analysis is heavily dependent on the B/H considerations. Previous iterations of modified B/H considerations have been applied inconsistently in MOA/HR analyses, which may be attributable in large measure to the availability of only relatively general, early guidance in this area (USEPA, 2005b; Sonich-Mullin *et al*., 2001). Some of the considerations have been misinterpreted due to a lack of common understanding of their appropriate level of application to MOA data in a WOE context; i.e., in overall data synthesis and evaluation of sufficiency of evidence to support a MOA decision versus the initial phase of systematic review (i.e., data selection and individual study review). Table 2 summarizes the variation in definitions of the B/H considerations in MOA analysis, which may also have contributed to inconsistency in application.

Evolved B/H considerations have been proposed and clarified here through delineation of the specific aspects addressed by each, as framed by a series of questions (captured below and summarized in Table 3). These questions build on those presented in Meek *et al*. (2014), based on additional experience in considering transparency in existing assessments as a basis to document comparative WOE. These evolved B/H considerations are proposed, then, not only as a basis to increase consistency in making WOE determinations for hypothesized MOA(s), but also to

Table 1. Template for weight of evidence based on evolved Bradford Hill considerations

Evolved Bradford Hill Considerations		Supporting Data	Inconsistent Data	Missing Data
1. **Biological Concordance**				
2. **Essentiality of Key events**				
3. **Concordance of Empirical Observations among Key Events**	Dose–response Temporality Incidence			
4. **Consistency**				
5. **Analogy**				

For a postulated mode of action, supporting data, inconsistent data and missing data are tabulated in the context of the evolved Bradford Hill considerations. Input in the supporting and inconsistent columns captures only what has been observed. Input in the missing column includes only that which is technically feasible and that is important for informing the mode of action. Cells are left blank in instances where data do not exist or are inadequate for evaluation. A brief narrative should accompany this table to describe the overall determination as to whether the data support or refute the hypothesis.

Table 2. Definition of the Bradford Hill considerations for application in mode of action analysis

Bradford Hill Considerations (Hill, 1965)	IPCS MOA/HR Framework (Boobis *et al.*, 2006; 2008; Sonich-Mullin *et al.*, 2001)	EPA Cancer Guidelines (USEPA, 2005b)	Evolved Bradford Hill considerations
Strength Strength of the association between suspected cause and observation.	**Strength** Unclearly defined. Considered together with specificity and consistency.	**Strength** The finding of large risks increases confi dence the association is not due to chance.	**N/A** Not considered applicable for evaluating MOA data.
Consistency Repeatability of an association by different persons, in different places, circumstances and times.	**Consistency** Repeatability of the key events in different studies. Considered together with strength and specificity.	**Consistency** Pattern of elevated risk observed across several independent studies.	**Consistency** Is the pattern of effects across species/ strains/ organs/test systems what would be expected?
Specificity The association is limited to a specific population and to particular sites and types of disease.	**Specificity** Stop/recovery studies show an absence or reduction of toxicity when a key event is blocked or reduced. Considered together with strength and consistency.	**Specificity** One cause associated with a single effect or disease.	**Essentiality of key events** Is the sequence of events reversible if dosing is stopped or a key event prevented?
Temporality The exposure occurs before the effect.	**Temporal association** Key events should be observable before toxicity is apparent.	**Temporal relationship** When exposure is known to precede development of the disease.	**Temporal concordance** Are the key events observed in hypothesized order?
Biological gradient Risk of disease increases with increasing exposure.	**Dose–response relationship** The dose–response for key events parallel the dose–response for the toxic effect. Increases in incidence of a key event correlate with increase in incidence of later key events.	**Biological gradient** Increasing effects associated with greater exposure.	**Dose–response concordance** Are the key events observed at doses below or similar to those associated with the end (adverse) effect?
Plausibility Biological knowledge supports suspected causation.	**Biological plausibility and coherence** Consistent with current understanding of biology. Considered together with coherence.	**Biological plausibility** Consistency with data from experiments or other sources demonstrating biological plausibility.	**Biological concordance** Does the hypothesized MOA conflict with broader biological knowledge? How well established is the MOA in the wider biological database?
Coherence The association agrees with the generally known facts of the history and biology of the disease.	**Coherence** Consistency with what is known specifically about the overall biological effects of the substance. Considered together with biological plausibility.	**Coherence** Information supporting cause and effect from other lines of evidence (i.e., animal bioassays, toxicokinetic studies and short-term studies).	**N/A** Not considered applicable for evaluating MOA data
Experiment Experimental evidence alters the frequency of associated events.	**N/A** Has not been mentioned in recent publications on the MOA/HR framework.	**Experimental evidence** when a change of exposure in a human popu lation brings about a change in disease.	**N/A** Not considered applicable for evaluating MOA data.

(Continues)

J. Appl. Toxicol. 2014; **34**: 595–606

wileyonlinelibrary.com/journal/jat

Journal of Applied Toxicology

Table 2. *(Continued)*

Bradford Hill Considerations (Hill, 1965)	IPCS MOA/HR Framework (Boobis *et al.*, 2006; 2008; Sonich-Mullin *et al.*, 2001)	EPA Cancer Guidelines (USEPA, 2005b)	Evolved Bradford Hill considerations
Analogy Information for a similar but different association supports causation. **N/A**	**N/A** Has not been mentioned in recent publications on the HR/MOA framework. **N/A** Considered as part of dose–response relationship definition.	**Analogy** insight gained from structure activity relationships and information on structural analogues. **N/A**	**Analogy** Would the MOA be anticipated based on broader chemical specific knowledge? **Incidence concordance** Is the occurrence of the end (adverse) effect less than that for preceding key events?

HR, human relevance; MOA, mode of action.

promote consistency in their application based on accumulating experience internationally.

The evolved B/H considerations are described in more detail below. These considerations appear in rank order based on their appropriate weighting of relative contribution to WOE determinations for hypothesized MOA(s), with those listed first contributing most significantly. Examples for evaluating weak to strong evidence for each evolved B/H consideration are also discussed.

Biological Concordance

- Does the hypothesized MOA conflict with broader biological knowledge?
- How well established is the MOA?

Evidence for a hypothesized MOA must satisfy the consideration of biological concordance. If available data on the hypothesized MOA are at odds with biological understanding, the hypothesis does not constitute a reasonable option for consideration. For instance, if a hypothesized early key event cannot conceivably lead to a subsequent hypothesized key or end event, it need not be considered.

The extent of evidence for biological concordance would be considered stronger, for example, if the hypothesized MOA has been well documented for a broad range of chemicals, and weaker if the hypothesized MOA is conceivable based on limited data or it has been hypothesized based simply on the possibility that none of the key events are at odds with biological understanding.

Essentiality of Key Events

- Is the sequence of events reversible if dosing is stopped or a key event prevented (i.e., counterfactual evidence)?

The extent of counterfactual evidence (i.e., experimental support for the necessity of a key event) is one of the principal determinants of WOE for a hypothesized MOA (Borgert *et al.*, 2011). For example, experimental evidence in animal models that lack a key metabolic pathway (e.g., knock out animal models) and fail to develop the end (adverse) effect would support essentiality of a key event. Similarly, if following cessation of repeated exposure for various periods, effects are reversible (i.e., late key events and/or the end (adverse) effect is prevented), this constitutes relatively strong evidence that key events are causal.

It is important to note that by its nature, counterfactual evidence typically addresses the necessity of an individual key event in a hypothesized MOA. Therefore, it may not always be helpful for discerning between two possible MOAs that share a key event. For example, if a chemical requires metabolic activation to be carcinogenic, a negative result in a 2-year cancer bioassay in an animal model null for the necessary activating enzyme supports that metabolism is necessary for carcinogenesis but is not helpful for differentiating between a MOA involving metabolic activation followed by direct DNA damage versus a MOA involving metabolic activation followed by cytotoxicity and regenerative proliferation.

Support for the essentiality of key events is considered stronger when there is direct counterfactual evidence supporting multiple key events in the hypothesized MOA. Evidence is considered weaker when evidence involves indirect measures for key events (i.e., the key event is inferred from the actual measured endpoint)

Table 3. Proposed changes to the Bradford Hill considerations and guidance for interpretation to improve application in the MOA/HR framework[a]

Evolved Bradford Hill considerations	Defining questions	Evidence for evaluating degree of support for the mode of action	
		Stronger	Weaker
1. Biological Concordance (replaces biological plausibility & coherence)	Does the hypothesized MOA conflict with broader biological knowledge? How well established is the MOA?	MOA is well established in scientific knowledge and/or completely consistent with established biological understanding.	MOA is contrary to well established biological understanding. MOA requires biological processes that are novel or poorly established.
2. Essentiality of Key Events (replaces strength, and specificity)	Is the sequence of events reversible if dosing is stopped or a key event prevented?	Counterfactual evidence to support key events (e.g., absence/reduction of later events when an earlier key event is blocked or diminished).	Data on reversibility only, indirect evidence only for key events or limited data available to assess.
3. Concordance of Empirical Observations among Key events (encompasses dose response and temporal concordance and beyond)	Dose–response: Are the key events observed at doses below or similar to those associated with end (adverse) effect? Temporality: Are the key events observed in hypothesized order? Incidence: Is the occurrence of the end (adverse) effect less than that for the preceding key events?	Dose–response and temporality: expected pattern of temporal and dose–response relationships based on robust database (multiple studies with examination of key events at interim time periods and at least 3 doses). Incidence: incidence of early key events is greater than end (adverse) effect.	All key events occur at all dose levels and all time points and/or limited data available to assess (e.g., inadequate dose spacing, missing key time periods for effect development, or failure to assess incidence at early time points). Incidence of early key events is lower than the end (adverse) effect and/or limited data available to assess.
4. Consistency (among different biological contexts)	Is the pattern of observations across species/strains/organs/test systems what would be expected based on the hypothesized MOA?	Pattern of effects are what would be expected across species, strains, organs and/or test systems.	Significantly inconsistent pattern of effects or limited data available to assess (e.g., effect only observed in a single rat strain).
5. Analogy (consistency across chemicals)	Would the MOA be anticipated based on broader chemical specific knowledge (e.g., the chemical is a member of a category for which related chemicals have known or strongly suspected MOA)?	Observations are consistent with those for other (related) chemicals having well defined MOA.	Pattern of effects for other (related) chemicals is distinctly different. Insufficient data to evaluate whether chemical behaves like related chemicals with similar proposed MOA.

MOA, mode of action.

[a]Evolution of the Bradford Hill (B/H) considerations for improved fit-for-purpose in the evaluation of sufficiency of data to support a hypothesized MOA. The evolved B/H considerations are rank ordered based on their appropriate weighting of relative contribution to weight of evidence determinations for hypothesized MOA(s), with those listed at the top contributing most significantly.

or non-specific inhibition of key events. For example, for a MOA hypothesized to involve binding to a receptor, demonstrating an end (adverse) effect is prevented by knocking-out or downregulating expression of the receptor is stronger than counterfactual evidence using a non-specific inhibitor.

Concordance of Empirical Observation Among Key Events

Concordance of empirical observations contributes considerably to the WOE for hypothesized MOA(s). Specifically, concordance of dose–response, temporality and incidence are key considerations. Each of these is addressed separately below. While not weighted as heavily as biological concordance and essentiality of key events, concordance of empirical observation across dose–response, temporality and incidence contributes significantly to WOE. Relationships and outliers should be carefully evaluated to understand whether the WOE strongly supports or is discordant with the hypothesized MOA, including consideration of cohesiveness across all three aspects of empirical observation.

Concordance of Dose–response Relationships Among Key Events

- Are the key events observed at doses below or similar to those associated with the end (adverse) effect?

In past MOA analyses, assessment of dose–response has sometimes been misinterpreted as simply addressing the question: "Is there evidence of a dose–response relationship for key events and/or the end (adverse) effect?" While this question is relevant to hazard characterization, it does not address dose–response concordance in relation to the WOE for a hypothesized MOA. Rather, the latter addresses the consistency of observed dose–response relationships among key and end (adverse) effects, as framed explicitly in the question above.

The hypothesized MOA is not supported in scenarios for which there is evidence that early key events occur only at higher doses than the end (adverse) effect. For example, a hypothesized receptor-based MOA is not supported by evidence indicating that receptor binding occurs only at doses well above those that cause frank liver injury, though it is important to consider if this might be a function of dose spacing in the relevant studies. Benchmark dose analyses for the dose–response relationships in key and end events are the most appropriate measure for consideration of their concordance, as they provide for direct comparison of comparable doses associated with a specified increase in each of the key events and/or end (adverse) effects and normalize for variations in dose spacing and group sizes in different studies.

Examination of the pattern of dose–response relationships is particularly important in considering the degree of support for hypothesized mutagenic MOAs (i.e., where mutation is an early and influential key event). For example, observation of a mutagenic response at high (cytotoxic) doses in genotoxicity assays is supportive of hypothesized MOAs where mutation is a secondary consequence of increased proliferative response resulting from tissue damage.

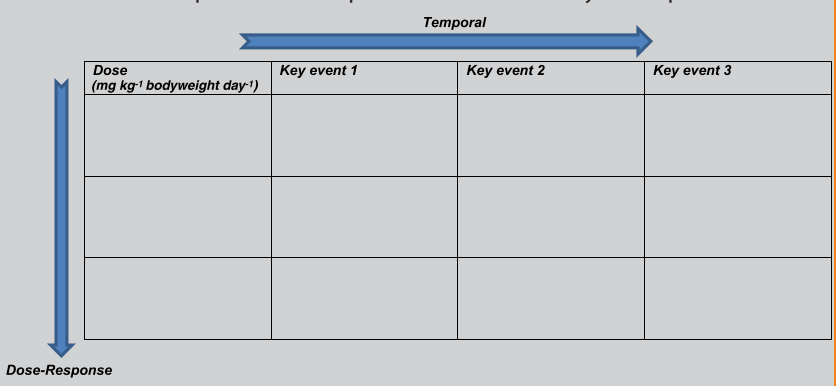
Concordance of Temporality (Time) Among Key Events

- Are the key events observed in hypothesized order?

Temporal concordance refers to the observation of key events in sequential order as described in the hypothesized MOA. In other words, earlier key events should be observed to precede later key events and the late (adverse) effect. Stronger evidence for temporal concordance is obtained when key events at interim time points demonstrate the hypothesized order (either in a single robust study or across multiple studies). Such evidence can often be acquired in studies examining the reversibility of key events and end (adverse) effects following various periods of exposure. Weaker evidence occurs when temporal data on key events are missing.

The template presented in Table 4 is often helpful in determining the extent to which evidence fulfills consideration of dose–response and temporal concordance in WOE analysis for MOA. If the hypothesized MOA is supported, the table should fill diagonally from the top left-hand corner to the bottom right-hand corner. This "pattern" supports a continuum of the relationship between early key events occurring at lower doses than late key events and outcome. Evidence of dose–response and temporal concordance is, for example, weaker if all key events occur at all dose levels and time points. Evidence is stronger, for example, if there is a reasonable range of studies of different durations with a minimum of three dose levels each and the "pattern" of results in this table (Table 4) is as described above.

Table 4. Dose–response and temporal concordance analysis template



Source: Meek and Klaunig (2010).

Journal of AppliedToxicology

Concordance of Incidence Between Key Events and End (Adverse) Effects

- Is the occurrence of the end (adverse) effect less than that for the preceding key events?

Clear evidence of the concordance of the incidence of the end (adverse) effect with that for early hypothesized key events is influential in contributing to WOE for hypothesized MOA(s). The incidence of hypothesized early key events should be greater than that for later key events and the (adverse) outcome, consistent with the important biological underpinning that key events are essential but not necessarily sufficient, to induce the relevant end (adverse) effect. For example, the hypothesis that cytotoxicity followed by regenerative proliferation are key events in the induction of specific tumors would be supported by the observation that the incidence of the former (cytotoxicity/regenerative proliferation) is greater than that for the latter (tumors) at a similar dose. "Incidence" here refers to the occurrence of a lesion of defined severity for each of the key and end events. It should be noted that a 1: 1 correlation of the incidence of early and late key events is not anticipated; lack of evidence for a 1: 1 correlation does not detract from contribution to the overall WOE. Consistent with the essentiality (but not necessarily sufficiency) of key events, lack of 1: 1 concordance is not unexpected, being a function of biological variability; i.e., lesions will not have progressed to the end (adverse) effect in all animals at the termination of exposure.

Consistency

- Is the pattern of observations across species/strains/organs/test systems what would be expected based on the hypothesized MOA?

Evidence of internal consistency within the collective data set for a chemical contributes to increased confidence in the WOE supporting a MOA. For example, if the initial hypothesized key event is oxidative metabolism to a reactive intermediate, are the target tissues and organs those which would be expected based on knowledge of distribution of the relevant metabolic enzyme? Evidence of consistency is stronger if the pattern of species-, strain- and sex-related variations in response is what would be expected based on known differences in metabolic profiles (e.g., extent and rate of metabolism to the putatively toxic entity). Evidence is weaker if there is either significant inconsistency in the expected pattern of the collective data based on the hypothesized MOA (e.g.,. the effect or result is only demonstrated in a single rat strain when data are available for multiple strains, for all of whom metabolizing capacity for the relevant pathway is anticipated to be similar) or when there are limited data available to assess this aspect.

Analogy

- Would the MOA be anticipated based on broader chemical specific knowledge?

Convincing evidence that the hypothesized MOA is operative for a broad range of chemically similar substances also contributes significantly to WOE. For example, consider the case where reductive metabolism for chemically similar substances is associated with a particular pattern of observations leading to the end (adverse) effect. If the pattern of observations for a related chemical is distinctly different, the evidence is weaker that these effects are produced by a similar MOA. On the other hand, if there is an extensive database illustrating that the MOA of interest is operative and leads to similar end (adverse) effects for several closely structurally related chemicals as identified, for example, by (quantitative) structure–activity modeling, evidence is stronger.

The rank order of the B/H considerations suggested above reflects their relative contribution to WOE determinations of MOA and is based on evolving experience internationally. In essence, data that conflicts with a broader biological understanding ranked highly here may be grounds for considering the available supporting data as inconsistent with the hypothesized MOA, whereas lack of concordance of some empirical data is often due to variations in, for example, dose spacing or administered doses in various studies and based on careful evaluation, would not detract meaningfully from the supporting database. In assessing the totality of the WOE, it is helpful to systematically take into account all of the considerations presented here as a basis to contribute to transparency in decision making. Such assessment benefits most from multidisciplinary input from both the relevant research and risk assessment communities. However, there is no minimum number of these evolved B/H considerations that must be met to determine sufficiency and/or associated confidence but rather, in their careful, systematic, more transparent and consistent consideration, cohesiveness (or not) of the supporting data becomes evident. It is also important to recognize that while some of the evolved B/H considerations may address the association of just one key event to the end event (e.g., essentiality of key events) the WOE determination is based on consideration of the interdependence of the key and end events in the hypothesized MOA.

Comparative Weight of Evidence Case Studies

To illustrate the utility of the comparative WOE approach, assessments for two chemicals (USEPA, 2009, 2010) were selected as case studies (i.e., carbon tetrachloride and TCP). The assessment of carbon tetrachloride drew on a previous evaluation of the US EPA (Manibusan *et al*., 2007), though the conclusions varied. These assessments were chosen based on the condition that B/H considerations for WOE had been explicitly addressed, consistent with the analysis in the MOA/HR framework for several potential MOA(s) for carcinogenicity. The focus here was not on the conclusions of the assessments; rather, the extensive review and synthesis of data therein provided the opportunity to address the potential utility of comparative analysis based on the evolved B/H considerations for WOE in MOA/HR analysis. As such, the evidence and conclusions were not re-evaluated but were simply extracted from the referenced assessments and summarized in the narrative tables presented (Tables 5a,b and 6) for the purpose of illustrating the methodology. Similarly, assessment of the underlying investigations was not considered, though based on the approach presented here, this might constitute an important next step. The literature reviews were also not updated, as the current analysis does not focus on particular chemicals but rather the potential value of the proposed methodology.

Carbon Tetrachloride

This analysis is based on a published hazard and dose–response assessment for carbon tetrachloride (USEPA, 2010). Carbon

Table 5. (a) Comparative weight of evidence analysis for carbon tetrachloride: cytotoxic MOA[a]

Evolved Bradford Hill considerations		Supporting data	Inconsistent data	Missing data
1. Biological concordance		Sustained cytotoxicity and proliferation is a well-established MOA for chemically mediated carcinogenicity.		
2. Essentiality of key events		No carbon tetrachloride induced liver toxicity in CYP2E1 knockout mice. CYP450 inhibitors prevent carbon tetrachloride liver damage. Mice treated with CYP450 inducers have increased carbon tetrachloride toxicity in subchronic and chronic studies.		
3. Concordance of empirical observations	Dose–response	Cytotoxicty and proliferation are observed at doses equal to or lower than doses at which tumors develop in rats and male mice	Tumors elevated at the lowest dose tested in female mice (5 ppm) without hepatocellular damage.	
	Temporality	Progression from cytotoxicity to hepatocellular proliferation is supported in acute and subchronic studies in rodents. Temporal relationship of cytotoxicity, repair, proliferation and tumor development is also supported in chronic cancer bioassay in rats.		Temporal relationship in female mice is not clearly defined.
	Incidence			
4. Consistency		Hepatic toxicity, necrosis and regenerative proliferation have generally been reported in animals exposed to carbon tetrachloride orally or by inhalation and are correlated with CYP450 content. Some evidence of DNA damage observed in concert with cytotoxicity.	One study reported development of tumors in mice at doses that did not produce necrosis but design of study may have influenced this result as animals were killed 1 month after last treatment.	
5. Analogy				
MOA, mode of action.				
[a]All conclusions in the above tables were extracted from the original US EPA toxicology review on carbon tetrachloride (USEPA, 2010).				
(b) Comparative weight of evidence analysis for carbon tetrachloride: mutagenic MOA[a]				
1. Biological concordance		Genotoxic MOA is well established for chemically mediated carcinogenicity.		
2. Essentiality of key events				
3. Concordance of empirical observations	Dose–response		Genotoxicity generally found at doses with cytotoxic effects.	Measurement of genetic damage to DNA has not been well

Journal of AppliedToxicology

		characterized at dose levels that do not cause cytotoxicity.
Temporality	Temporality not observed. Genotoxicity generally found in concert with cytotoxicity.	
Incidence	Extensive *in vitro* and *in vivo* genotoxic data are primarily negative.	Genetically damaging events occurring at or below doses that induce cytotoxicity in laboratory rodents.
4. Consistency	Doses where cytotoxic events are observed are lower than doses for which mutagenicity has been evaluated.	
	Limited positive results in genotoxicity assays appear more related to a cytotoxic response than to a mutation event	
5. Analogy		

MOA, mode of action.
[a]All conclusions in the above tables were extracted from the original US EPA toxicology review on carbon tetrachloride (USEPA, 2010).

tetrachloride caused hepatocellular adenomas and carcinomas in rats, mice and hamsters in oral studies and in rats and mice following inhalation exposure. In addition to liver tumors, adrenal pheochromocytomas were observed in male and female mice following oral and inhalation exposure, for which it was concluded that data were inadequate to evaluate MOA. There was no increase in pheochromoctyomas in rats.

Based on the analysis of available data, including that on MOA, it was concluded in the assessment (USEPA, 2010) that the agent is likely a human carcinogen. Further, a potential MOA for carbon tetrachloride-induced liver tumors was hypothesized, with the following key events that included: (1) metabolism to the trichloromethyl radical by CYP2E1 and subsequent formation of the trichloromethylperoxy radical; (2) radical-induced damage leading to hepatocellular toxicity; and (3) sustained regenerative and proliferative changes in the liver in response to hepatotoxicity. The possibility that carbon tetrachloride may act via a mutagenic MOA (i.e., where mutation is an influential early key event in the induction of tumous versus, for example, being secondary to tissue damage) was also considered but not evaluated in a manner based on WOE considerations consistent with the MOA/HR framework. Based on the inconsistencies in the database supporting a potential role for the cytotoxicity, regenerative, proliferation-based MOA at the low end of the experimental exposure range and the complexity of the genotoxicity database, it was concluded that, "… the carcinogenic MOA for carbon tetrachloride is not known. Therefore, consistent with the *Guidelines for Carcinogen Risk Assessment* (USEPA, 2005b), linear low-dose extrapolation as a default approach was applied to data for liver tumors and pheochromocytomas" (USEPA, 2010).

1,2,3-Trichloropropane

This analysis is based on a hazard and dose–response assessment of TCP released in 2009 (USEPA, 2009). Based on the observed statistically significant dose-related increases in multiple tumor types in both sexes of rats and mice in a 2-year carcinogenicity assessment (NTP, 1993) and related mechanistic data (including that on genotoxicity), it was concluded that TCP is "likely to be carcinogenic to humans" via a mutagenic MOA. Relevant data for alternative MOA(s) such as cytotoxicity with tissue repair and disruption of cell signaling were considered insufficient to evaluate. It was further concluded that the available data support a hypothesized mutagenic MOA with two key events: (1) metabolism to a DNA-reactive compound, and (2) (early) induction of mutations. A low-dose linear extrapolation approach to dose–response analysis was applied, consistent with the *Guidelines for Carcinogen Risk Assessment* (USEPA, 2005b)

Comparative Weight of Evidence Analysis

Narrative comparative WOE summary tables were constructed for the hypothesized and alternative MOA(s) for carbon tetrachloride (Table 5a,b) and for a mutagenic MOA for TCP (Table 6) based on the consideration and evaluation of the data in the existing assessments (USEPA, 2009, 2010). For each postulated MOA, supporting data, inconsistent data and missing information were tabulated in the context of the evolved B/H considerations. As per MOA/HR framework recommendations, the information in the supporting and inconsistent data columns capture what has been observed, not what might be possible if more experiments had been performed. In addition, the

Table 6. Comparative weight of evidence analysis for 1,2,3-trichloropropane: mutagenic MOA

Evolved Bradford Hill considerations		Supporting data[a]	Inconsistent data[a]	Missing data[b]
1. Biological concordance		Genotoxic MOA is well established for chemically mediated carcinogenicity		
2. Essentiality of key events		Inducers/inhibitors of metabolism alter amount of DNA binding		Evidence for adduct conversion to genetic damage
3. Concordance of empirical observation	Dose–response	Dose-related formation of DNA-reactive metabolite, DNA adduct formation, tumor formation and time to tumor.		
	Temporality	Metabolism to reactive intermediate occurs within hours of exposure, adducts appear within hours and days of exposure, and tumors first appear after ≈ 9 months.		
	Incidence			No data to assess whether adduct formation frequency different from tumor frequency.
4. Consistency		Mutagenic effects *in vitro* accompanied by limited evidence of *in vivo* mutagenicity.	Adducts occur in tissues where no neoplastic effects were reported (spleen, liver and glandular stomach). Negative results from *in vivo* genotoxicity assessments (dominant lethal and micronucleus).	
5. Analogy		Other halogenated aliphatic chemicals (1,2,-dibromoethane and 1,2-dibromo-3-chloropropane) are mutagenic carcinogens. Other genotoxic chemicals are multisite and multispecies carcinogens.		

MOA, mode of action.

[a]All conclusions in the above tables were extracted from the original US EPA toxicology review on 1,2,3-trichloropropane (USEPA, 2009).

[b]The IRIS assessment did not comment on missing data; the information here represents the authors' views.

Journal of AppliedToxicology

information noted in the missing column only includes that which is testable and important for informing the MOA (i.e., critical data needs). Ideally, a discussion on whether the missing information is critical and would detract from or impact conclusions regarding the proposed MOA should accompany this comparative WOE table. Blank cells would typically represent instances where data either do not exist or are inadequate for evaluation. However, in this case, as the analysis draws upon an existing assessment, blank cells may also represent where text was either absent or inadequate to address the evolved B/H considerations.

Qualitative Assessment of Overall Evidence

For both case studies, the focus is not to conclude on the sufficiency of underlying data to support a particular MOA conclusion, but rather to illustrate the utility of the comparative WOE approach for increasing transparency in the assimilation of data.

Visually, Tables 5(a,b) and 6 highlight the availability of supporting and discrepant data on the MOA(s) evaluated for carbon tetrachloride and TCP. Comparative WOE analysis, for the two hypothesized MOA(s) for carbon tetrachloride based on the published assessment (USEPA, 2010), indicates that the supporting data for the hypothesized MOA involving cytotoxicity (necessarily within the range of experimental observation) fulfill a number of the evolved B/H considerations. This contrasts with the comparatively more limited support for the hypothesized mutagenic MOA. This difference highlights:

(1) the potential utility of comparative analysis for assessing the WOE of alternative MOA(s) for individual chemicals, based on the evolved B/H considerations to more explicitly indicate the degree of confidence in a particular MOA, and
(2) the desirability, in the interest of transparency and consistency, of separating conclusions reflecting assessment of the relative WOE for MOA in the observable experimental range based on articulated and explicit considerations from those based on inference or extrapolation to the low-dose range. It is anticipated that such an approach has the potential to increase transparency in delineating science judgment determinations from those related to public policy.

The comparative WOE analysis for TCP also provides a basis for comparison across chemicals of a relatively strong database for a mutagenic MOA, which can be contrasted with one that is relatively weak, potentially as a basis to increase consistency in determinations. In this case, perspective on the degree of confidence in the supporting WOE for the hypothesized mutagenic MOA for carbon tetrachloride (Table 5b) can be gained through comparison with the nature and extent of data available for the stronger database for TCP (Table 6).

Discussion

Comparative aspects of WOE analyses are illustrated here as a basis to contribute to transparency and consistency in delineating confidence/uncertainty in MOA/HR analysis based on the BH considerations. As noted by Guyton *et al*. (2008), Hill's (1965) considerations were not developed originally for evaluation of experimental/mechanistic data, though their utility for application in modified form to assess WOE in MOA analysis has been repeatedly though inconsistently tested. Based on increasing experience internationally in MOA/HR analysis (see, for example, Boobis *et al*., 2006, 2008, Meek *et al*., 2014), evolved B/H considerations are proposed here and clarified through delineation of the specific aspects addressed by each as framed by a series of questions. Definitions for these considerations have been additionally simplified and tailored to application in MOA analysis. The evolved B/H considerations were also rank ordered to reflect their relative contribution to WOE determinations and their utility exemplified in a comparative WOE approach.

The evolved B/H considerations build on previously published iterations and reflect experience in the application of MOA analysis. Several terms were clarified to facilitate assimilation of relevant chemical specific and biological data (i.e., "specificity" is now termed "essentiality of key events," "biological plausibility and coherence" is now termed "biological concordance" and concordance of empirical observations among key events delineated). In addition, considerations with limited relevance for evaluating MOA data (i.e., "strength," "coherence" and "experiment") were eliminated while other considerations (i.e., "analogy" and "incidence concordance") were added based on evolving experience with larger numbers of chemicals. It is hoped this evolved terminology, which reflects more common understanding within the broader risk assessment (versus epidemiological) community, will additionally contribute to consistency of use in MOA analysis. Finally, considerations were redefined as a basis to promote consistency and utility. For example, in publications of the IPCS MOA/HR framework (Boobis *et al*., 2006, 2008; Sonich-Mullin *et al*., 2001), consistency is defined as repeatability of key events in different studies; while in the USEPA cancer guidelines, consistency refers to the pattern of elevated risk observed across several independent studies (USEPA, 2005b). Neither definition accurately reflects the use of consistency in evaluating the WOE for hypothesized MOA(s). The former simply assesses reproducibility of results and, as such, may only contribute to the level of confidence in the occurrence of one key event. The latter definition is more appropriate to the assessment of the reproducibility of results in epidemiological and not mechanistic data sets. Consistency in the context of the MOA/HR framework more appropriately relates to evaluation of the WOE supporting interdependence of the key and end (adverse) events. Therefore, consistency was redefined here to reflect support of the pattern of effects across species/strains/organs and test systems for the hypothesized MOA. For example, if metabolism is a hypothesized key event in a carcinogenic MOA, the pattern of species-, strain- and sex-related variations in tumor response is compared to that expected based on known differences in metabolic profiles in the test systems. As such, it is not as important to assess if the occurrence of tumors is reproducible across studies, but rather, if the presence or absence of tumors in various species and strains is consistent with the hypothesized MOA.

Comparative WOE analysis is illustrated as a means of increasing understanding of the nature of transparency that is essential when evaluating confidence in the supporting WOE for hypothesized (potentially competing) MOAs. In doing so, it also provides a basis for increasing consistency in evaluation. Presentation of an overview of the data in a comparative manner (i.e., as supporting, inconsistent and missing) based on templates that cue evaluators concerning critical aspects provides concise insight into the extent of available data and relevant patterns in the existing database, which support various levels of confidence in considered options. In addition, this presentation concisely communicates areas of uncertainty (inconsistent data column and blank cells) and highlights areas of greatest impact for future research (missing data column). Ideally, further transparency on

the impact of this information (i.e. supporting, inconsistent and missing data) on the MOA conclusions would be provided in a detailed, supplemental discussion.

Synthesis of a collective data set to evaluate WOE for a hypothesized MOA is complex and challenging, requiring multidisciplinary input from both the research and risk assessment communities. This analysis is dependent upon transparent and consistent evaluation of the extent and nature of both chemical-specific and biological data versus supposition about possibilities for which there is essentially no experimental support. Characterization of the evolved B/H considerations is anticipated to contribute to more robust and transparent analyses, as a basis also to discourage, without clear rationale, the discounting of well-supported options based on the emphasis of outlying data of lesser quality.

This manuscript extends MOA/HR assessment through evolution of the B/H considerations and illustration of a comparative WOE analysis. Ultimately, it is anticipated that the additionally articulated and comparative aspects, which build on considerable recent experience in MOA analysis, will contribute to increasing transparency, consistency and methodological rigor in separating aspects of science judgment (i.e., weighting of options based on transparent consideration of available scientific support) from those of public policy in regulatory risk assessment (the latter of which sometimes involves embedded conservatism, to increase public health protection).

Acknowledgements

This work builds on previous products of a significant number of contributing authors and collaborators who have participated in the development and application of mode of action/species concordance analysis and associated training materials. Contributors include those outlined in Meek *et al*. (2003, 2014), Seed *et al*. (2005) and Boobis *et al*. (2006, 2008).

Conflict of Interest

Several of the authors (C.M.P., A.N.B., C.M.N. and R.J.L) are employed by a subsidiary of Exxon Mobil, who produces materials evaluated by the US EPA. Methodological aspects based on case studies considered here do not relate to specific evaluations of relevance to Exxon Mobil.

The work reported in the paper was conducted during the normal course of research/training for the University of Ottawa (MEM). It evolved from discussions during and following provision of training, for which funding for development of materials, travel and lodging was provided.

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-Hirsche B, Meek ME, Munn S, Ruchirawat M, Schlatter J, Seed J, Vickers C. 2008. IPCS framework for analyzing the relevance of a non-cancer mode of action for humans. *Crit. Rev. Toxicol.* **38**: 87–96.

Borgert CJ, Mihairch 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.* **61**(2): 185–191.

Carmichael N, Bausen M, Boobis AR, Cohen SM, Embry M, Fruijtier-Polloth C, Greim H, Lewis RJ, Meek ME, Mellor H, Vickers C, Doe F. 2011. Using mode of action information to improve regulatory decision-making: an ECETOC/ILSI RF/HESI workshop overview. *Crit. Rev. Toxicol.* **41**: 175–186.

Dellarco VL, Baetcke K. 2005. A risk assessment perspective: application of mode of action and human relevance frameworks to the analysis of rodent tumor data. *Toxicol. Sci.* **86**: 1–3.

Guyton KZ, Barone S, Brown RC, Euling SY, Jinot J, Makris S. 2008. Mode of action frameworks: a critical analysis. *J. Toxicol. Environ. Health B* **11**: 16–31.

Hill AB. 1965. The environment and disease: An association of causation. *Proc. R. Soc. Med.* **58**: 295–300.

Manibusan MK, Odin M, Eastmond DA. 2007. Postulated carbon tetrachloride mode of action: a review. *J. Environ. Sci. Health.* **25**(3): 189–209.

Meek ME. 2008. Recent developments in frameworks to consider human relevance of hypothesized modes of action for tumours in animals. *Environ. Mol. Mutagen.* **49**(2): 110–116.

Meek ME, Klaunig J. 2010. Interpreting available data and identifying critical data gaps for benzene risk assessment. *Chem. Biol. Interact.* **184**, 279–285.

Meek ME, Bucher JR, Cohen SM, Dellarco V, Hill RN, Lehman-McKeeman LD, Longfellow DG, Pastoor T, Seed J, Patton DD. 2003. A framework for human relevance analysis of information on carcinogenic modes of action. *Crit. Rev. Toxicol.* **2003**(33): 591–653.

Meek ME, Berry C, Boobis AR, Cohen SM, Hartley M, Munn S, Olin SS, Schlatter J, Vickers C. 2008. Mode of action frameworks: a critical analysis. *J. Toxicol. Environ. Health B* **11**: 681–685.

Meek ME, Boobis A, Cote I, Dellarco V, Fotakis G, Munn S, Seed J, Vickers C. 2014. 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.

NTP. 1993. Toxicology and carcinogenesis studies of 1,2,3-trichloropropane (CAS No. 96-18-4) in F344/N/N rats and B6C3F1 mice (gavage studies). Public Health Service, U.S. Department of Health and Human Services; NTP TR 384. NIH Publication No. 94–2839.National Institute of Enviornmental Health Sciences: Research Triangle Park, NC.

SAB. 1999. Minutes of the Chloroform Risk Assessment Review Subcommittee Meeting, October 27–28, 1999.

SAB. 2007. Subject: Advisory on EPA's assessments of carcinogenic effects of organic and inorganic arsenic: a report of the US EPA Science Advisory Board. Available from: http://www.yosemite.epa.gov/sab/sabproduct.nsf/EPA-SAB-07-008.pdf.

SAP. 2000. Atrazine: Hazard and dose–response assessment and characterization, SAP Report No. 2000–05. FIFRA Scientific Advisory Panel Meeting, June 27–29, 2000. Available from: http://www.epa.gov/scipoly/sap/meetings/2000/june27/finalatrazine.pdf.

Seed J, Carney E, Corley ER, Crofton K, DeSesso J, Foster P, Kavlock R, Kimmel G, Klaunig J, Meeik E, Preston J, Slikker W, Tabacova S, Williams G. 2005. Overview: Using mode of action and life stage information to evaluate the human relevance of animal toxicity data. *Crit. Rev. Toxicol.* **35**: 663–672.

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 ME, Rice J, Younes M. 2001. IPCS conceptual framework for evaluating a mode of action for chemical carcinogenesis. *Regul. Toxicol. Pharmacol.* **34**: 146–152.

USEPA. 2005a. Science Issue Paper: Mode of carcinogenic action for cacodylic acid(dimethylarsinic acid, DMAV) and recommendations for dose response extrapolation July 26, 2005, Health Effects Division, Office of Pestcide Programs, US Environmental Protection Agency. Available from: http://www.epa.gov/oppsrrd1/reregistration/cacodylic_acid/dma_moa.pdf.

USEPA. 2005b. Guidelines for Carcinogen Risk Assessment, EPA/630/P-03/001 F. Risk Assessment Forum, US Environmental Protection Agency, Washington DC, March 2005.

USEPA. 2009. Toxicological Review of 1,2,3-Trichloropropane (CAS No. 96-18-4) in Support of Summary Information on the Integrated Risk Information System (IRIS), EPA 635/R-08/010 F, US Environmental Protection Agency, Washington DC, September 2009.

USEPA. 2010. Toxicological Review of Carbon Tetrachloride (CAS No. 56-23-5) in Support of Summary Information on the Integrated Risk Information System (IRIS), EPA 635/R-08/005F, US Environmental Protection Agency, Washington DC, March 2010.