

Evaluating the Human Relevance of Chemically Induced Animal Tumors

Samuel M. Cohen,^{*1} James Klaunig,[†] M. Elizabeth Meek,[‡] Richard N. Hill,[§] Timothy Pastoor,[¶] Lois Lehman-McKeeman,[|] John Bucher,^{||} David G. Longfellow,^{|||} Jennifer Seed,[§] Vicki Dellarco,[§] Penelope Fenner-Crisp,[#] and Dorothy Patton[#]

^{*}ILSI RSI Steering Committee, University of Nebraska Medical Center, Omaha, Nebraska 68198-3135; [†]Indiana University School of Medicine, Indianapolis, Indiana 46202-5120; [‡]Health Canada, Ottawa, Ontario K1A 0L2 Canada; [§]U.S. Environmental Protection Agency, Washington, District of Columbia 20460; [¶]Syngenta Crop Protection, Greensboro, North Carolina 27419-8300; [|]Bristol-Myers-Squibb Company, Princeton, New Jersey 08543; ^{||}National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina, 27709; ^{|||}National Cancer Institute, National Institutes of Health, Rockville, Maryland 20892-7638; and [#]International Life Sciences Institute (ILSI), Washington, District of Columbia 20005

Received December 23, 2003; accepted December 31, 2003

Defining the mode(s) of action by which chemicals induce tumors in laboratory animals has become a key to judgments about the relevance of such tumor data for human risk assessment. Frameworks for analyzing mode of action information appear in recent U.S. EPA and IPCS publications relating to cancer risk assessment. This FORUM paper emphasizes that mode of action analytical frameworks depend on both qualitative and quantitative evaluations of relevant data and information: (1) presenting key events in the animal mode of action, (2) developing a “concordance” table for side-by-side comparison of key events as defined in animal studies with comparable information from human systems, and (3) using data and information from mode of action analyses, as well as information on relative sensitivity and exposure, to make weight-of-evidence judgments about the relevance of animal tumors for human cancer assessments. The paper features a systematic analysis for using mode of action information from animal and human studies, based in part on case examples involving environmental chemicals and pharmaceuticals.

Key Words: carcinogenic mode of action; human relevance of animal tumors; risk assessment; PPAR α agonists.

For several decades, potential carcinogenic hazard to humans has been identified primarily on the basis of long-term animal bioassays. Relatively standardized protocols have been established for studies in rodents, particularly rats and mice. As in all experiments in animals, two fundamental assumptions are made: (1) the results in the animal bioassay are relevant to

humans (interspecies extrapolation); and (2) the doses used in the animal bioassay are relevant for estimating risk at known or expected human exposure levels (dose extrapolation). Although these assumptions are valid for many chemicals with respect to carcinogenesis, progress during the past four decades in our understanding of the mode of action of carcinogenesis has greatly improved our ability to rigorously evaluate these assumptions for specific chemicals. This is critical for a rational assessment of hazard and risk to humans.

A fundamental breakthrough in our progress in understanding mode of action was the distinction between DNA reactivity and non-DNA reactivity, more broadly referred to as genotoxic versus nongenotoxic modes of action. As progress has been made in determining the modes of action of chemicals that produce neoplasia in animal assays, it has become increasingly important to evaluate the relevance of these modes of action with respect to humans.

To address the issue of the human relevance of the mode(s) of action determined in animals, a working group was formed under the sponsorship of the U.S. Environmental Protection Agency (EPA) and Health Canada, organized by the International Life Sciences Institute Risk Science Institute (ILSI RSI). Based on the mode of action frameworks developed by the International Programme on Chemical Safety (IPCS) (Sonich-Mullin *et al.*, 2001) and the U.S. EPA (U.S. EPA, 1999), commonly used today by other regulatory agencies and international organizations (e.g., the World Health Organization, Expert Panel of the Joint Meeting on Pesticide Residues), RSI charged the working group with expanding these frameworks to include evaluating the human relevance of modes of action determined in animals. The details of the process, the case studies, and the framework are published as a series of papers in the November 2003 issue of *Critical Reviews in Toxicology (CRT)* (Cohen *et al.*, 2003; Meek *et al.*, 2003). The present article briefly describes the framework and provides a user's guide for its application. While brief references to specific

The views expressed in this report are those of the individual authors and not necessarily those of the organizations they represent.

Preliminary versions of this work were presented at the Society for Risk Analysis meeting in New Orleans in December 2002 and at the Society of Toxicology Meeting in Salt Lake City in March 2003.

¹To whom correspondence should be addressed at University of Nebraska Medical Center, Department of Pathology/Microbiology, 983135 Nebraska Medical Center, Omaha, NE. Fax: (402) 559-9297. E-mail: scohen@unmc.edu.

Toxicological Sciences vol. 78 no. 2 © Society of Toxicology 2004; all rights reserved.

examples on which the framework is based are included, the details are provided in the *CRT* papers mentioned above.

Several iterations of case studies of chemicals with generally well-known modes of action were used to develop this framework. The intent is to provide guidance for a disciplined, transparent process evaluating the mode of action in animals and each key event (see below) with respect to human relevance.

The expanded human relevance framework is based on three fundamental questions, followed by an explicit description of confidence in the evaluation, identification of specific data gaps, and the implications for risk assessment. It is emphasized that this framework is intended to be used as part of the hazard characterization step of the overall risk assessment process.

THE FRAMEWORK

The human relevance framework developed by the working group is summarized as follows:

1. Is the weight of evidence sufficient to establish the mode of action (MOA) in animals?
2. Are key events in the animal MOA plausible in humans?
3. Taking into account kinetic and dynamic factors, is the animal MOA plausible in humans?
4. Conclusion: Statement of confidence, analysis, and implications.

Site concordance between animals and humans is a fundamental initial premise on which mode of action frameworks are based. As a result, in applying this framework for a given chemical, tumors of each animal target organ observed are evaluated independently, with the assumption that different modes of action are possible in different organs, though based on this analysis, modes of action in different tissues may be similar. Similarly, an evaluation of the likelihood of congruence between target organ(s) in different species and in humans needs to be made, based on the mode of action analysis.

Is the Weight of Evidence Sufficient to Establish the Mode of Action in Animals?

For each tumor type identified in an animal bioassay, the current EPA and IPCS guidance frameworks (Sonich-Mullin, *et al.*, 2001; U.S. EPA, 1999) delineate specific topics for organizing and presenting the information. The approach is based on the Hill criteria for causality, originally developed for application in epidemiologic investigations (Hill, 1965). IPCS emphasized that its framework "is not a checklist of criteria, but rather presents an analytical approach to considering the weight-of-evidence of an MOA."

The process of evaluating a mode of action in animals is based on an explicitly stated proposed mode of action and enumeration of key events. As defined by IPCS and EPA, key events are measurable parameters associated with *critical* steps

in the mode of action. These steps and the framework for evaluating an animal MOA are as follows: postulated MOA; key events; associated critical parameters; dose-response relationships; temporal association; strength, consistency, and specificity of association of key events and tumor response; biological plausibility and coherence; possible alternative MOAs; conclusion about the MOA; uncertainties, inconsistencies, and data gaps.

This process incorporates an evaluation of the weight of evidence for possible alternative MOAs at a given site and an evaluation of the overall strength of evidence supporting the MOA under consideration. Ultimately, a decision concerning the weight of evidence supporting the MOA and the level of confidence in that decision must be made. It also identifies critically important data gaps that would increase confidence in potential modes of action.

For a given chemical, the primary sources of information for evaluating an MOA in animals often likely are data generated for that specific chemical in the animal model. Obviously, data from other sources can and should also be used, as appropriate, along with data on chemicals with similar chemical structures, similar modes of action, or both. If the mode of action for a chemical is novel, considerably more data will be required to support the conclusion that it is related to the carcinogenic process of the tumors induced by that chemical than subsequent examples of chemicals acting by the same mode of action. The working group did not address the issue of how much data is sufficient to support a specific mode of action for a given chemical, except by way of example within the case studies. Consideration at this stage of the mode of action in the context of potential variations between animals and humans also facilitates addressing subsequent steps in the framework.

Are Key Events in the Animal MOA Plausible in Humans?

This represents a qualitative assessment of the relevance of the MOA to human cancer potential. Listing the critical specific key events that occur in the animal mode of action and directly evaluating whether each of the key events might or might not occur in humans facilitated consideration and transparent presentation of the relevant information. Presentation in tabular form (as shown in Table 1 for atrazine), referred to as a concordance table, can be helpful in delineating the relevant information. The key events (and possibly some of the critical associated processes) are listed with the information regarding these events for the animal in which the tumor was observed. It is intended that the information in these tables be brief, since a narrative explanation is expected to accompany the table. In the right-hand column, the effect in humans for each of the key events is evaluated. An additional column for the results in a different strain, species, or route of administration that does not result in tumors can be useful if information is available for comparison to the model that leads to tumors.

The evaluation of the concordance for a given chemical in

TABLE 1
Concordance of Key Events in Rats and Humans: Atrazine

Key event	Evidence in animals (rats)		Evidence in humans
	Sprague-Dawley	F-344	
Decreased GnRH pulses	Yes	No	Unknown/possible
Suppression of LH surge	Yes	No	Unknown/possible
Change in cyclicity	Yes	No	No/not relevant
Prolonged increase in estrogen/prolactin	Yes	No	No

humans is an evaluation of the MOA in humans, rather than an evaluation of the specific chemical. What kinds of information can be utilized to evaluate the key events in humans? Essentially, such data can come from *in vitro* and *in vivo* studies on the substance itself, but also can involve basic information regarding anatomy, physiology, genetic disorders, epidemiology, and any other information that is known regarding the key events in humans. Information concerning an evaluation of the key event in humans exposed directly to the chemical is frequently unavailable.

In evaluating the concordance of the information in humans to that in animals, a narrative describing the weight of evidence and an evaluation of the level of confidence for the human information needs to be provided. Some specific types of information that are useful include the following: cancer incidences at the anatomical site and cell type of interest, including age, sex, ethnic differences and risk factors, including chemicals and other environmental agents; knowledge of the nature and function of the target site, including development, structure (gross and histologic), and control mechanisms at the physiological, cellular, and biochemical levels; human and animal disease states that provide insight concerning target organ regulation and responsiveness; human and animal responses to the chemical under review or analogs following short, intermediate, or long-term exposure, including target organs and effects.

Obviously, a substantial amount of information is required to conclude that the given mode of action is not relevant to humans. If such a conclusion is strongly supported by the data, then chemicals producing animal tumors only by that mode of action would not pose a cancer hazard to humans, and no additional risk assessment is required. Since there is no cancer hazard, there is no cancer risk for the tumor under consideration.

Taking Into Account Kinetic and Dynamic Factors, Is the Animal MOA Plausible in Humans?

Toxicokinetic and toxicodynamic data are customary and valuable contributions to the overall risk assessment process. For purposes of human relevance analysis, if the animal MOA

is judged to be qualitatively relevant to humans, a more quantitative assessment is required, taking into account any toxicokinetic and toxicodynamic information that is available in both the animal and in humans. Such data will, of necessity, be both chemical and mode of action specific, requiring information regarding the chemical and the mode of action in animals and humans for comparison. Toxicokinetic considerations include nature and time course of chemical uptake, distribution, metabolism, and excretion, whereas toxicodynamic considerations include the consequences of the interaction of the chemical with cells. Of course, consideration of physiological, cellular, and biochemical differences between species regarding endogenous chemicals and control systems may be required.

Like the qualitative assessment, a tabular comparison between the animal and humans regarding quantitative issues can be helpful. Similarly, the data for the quantitative assessment in humans may not directly involve exposure to the chemical under consideration, but may incorporate a more general evaluation in humans, involving information on the key events and biologic processes that are known in general, rather than limited specifically to the chemical being evaluated. However, this is the quantitative comparison, so it is a bit more of a stretch to use "other data"—more so than for qualitative modes of action.

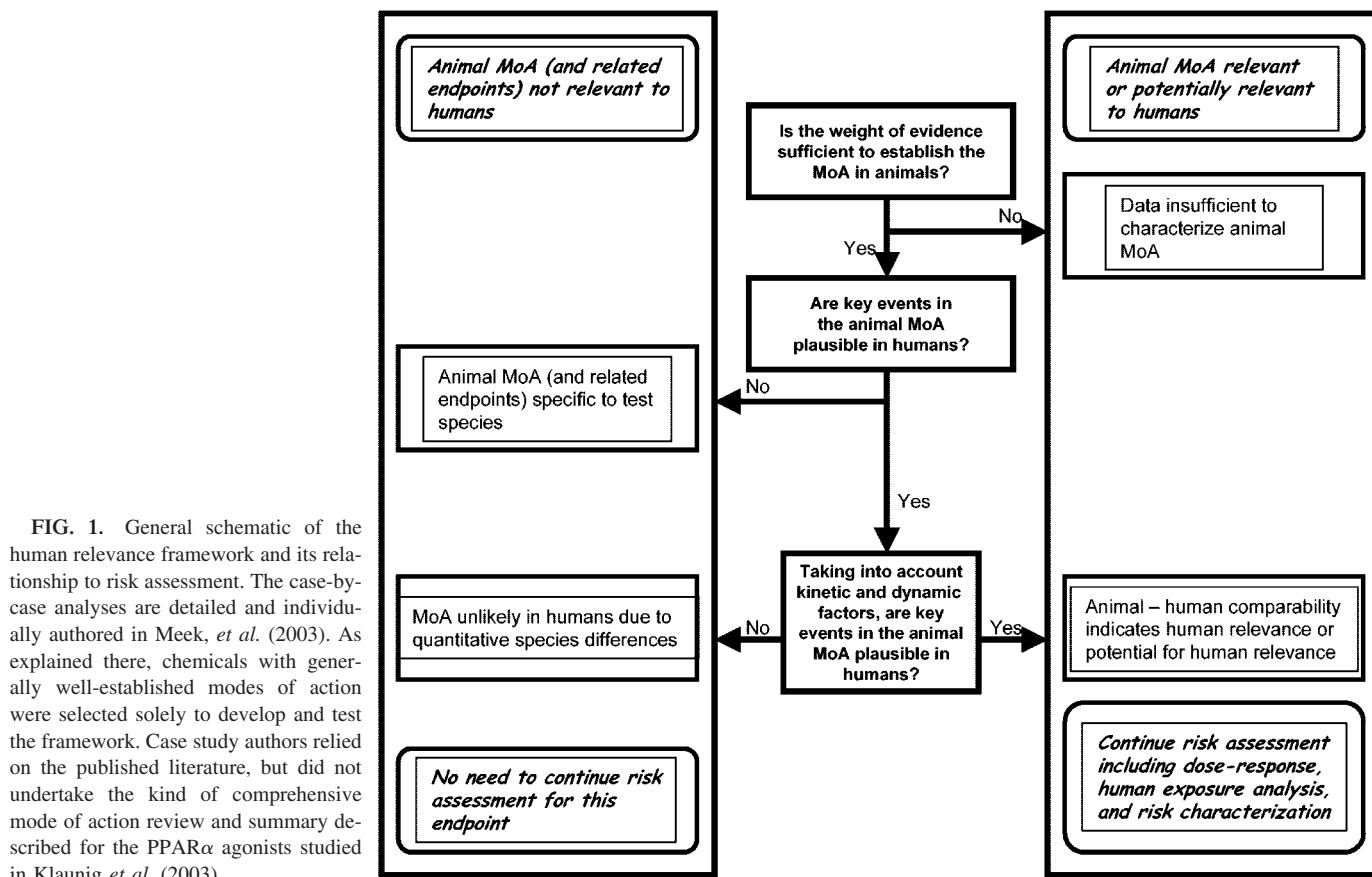
Statement of Confidence, Analysis, and Implications

Following the overall assessment of each of the three questions of the human relevance framework, a statement of confidence is necessary that addresses the quality and quantity of data underlying the analysis, consistency of the analysis within the framework, consistency of the database, and the nature and extent of the concordance analysis. An evaluation of alternative modes of action, using comparable analyses and rigor, is also essential. A critically important outcome of adequate consideration of the weight of the evidence for an overall mode of action and the qualitative and quantitative concordance is the identification of specific data gaps that can be addressed experimentally in future investigations to increase confidence.

APPLICATION OF THE FRAMEWORK: SOME EXAMPLES

An overall general schematic of the human relevance framework and its relationship to risk assessment is shown in Figure 1. It is essential that, even if a specific mode of action cannot be identified for a given chemical or if the mode of action is qualitatively and quantitatively applicable to humans, the specific information evaluated during this analysis should be incorporated into the remaining steps of the risk assessment. In many instances, critical pieces of information needed for evaluation of human risk will have been identified in this framework analysis.

The modes of actions and chemicals that were evaluated by



the working group included direct or indirect genotoxicity (acrylonitrile), interaction with proteins (*d*-limonene), and suppression of luteinizing hormone (atrazine), increased hepatic clearance of thyroxin (phenobarbital), sustained cytotoxicity and regeneration (chloroform), and urinary tract calculi (melamine). In addition, ethylene oxide was evaluated as an example of a DNA-reactive carcinogen.

Two modes of action have been postulated for the carcinogenicity of acrylonitrile in animal models, particularly related to brain tumors: either direct DNA-reactivity of acrylonitrile metabolites, or indirect DNA-reactivity by induction of oxidative damage. It was concluded by our working group that the data were insufficient to describe definitively the mode(s) of action regarding acrylonitrile, so the answer to question 1, "Is the weight of evidence sufficient to establish the mode of action in animals?" was "no." The consequence of this decision is that the default assumption that the mode of action in the animal bioassay is relevant to humans applies.

In contrast, the modes of action of the other chemicals used for case studies have been well delineated. *d*-Limonene produces kidney tumors in male rats, but not in female rats or either sex of mice. This has been attributed to the binding of the epoxide of *d*-limonene to α 2u-globulin, which is absorbed into the proximal tubule but not digested, leading to toxicity of the tubular cells and consequent cellular regeneration and

ultimately tumor formation. Although humans can metabolize *d*-limonene to its epoxide, there is no functionally similar protein in humans to which the epoxide can bind. Thus, one of the critical events in the process, binding to α 2u-globulin or a similar protein, is precluded. Therefore, it can be concluded that this MOA does not occur in humans. As a consequence, while the answer to the question 1 was "yes," the answer to question 2, "Are key events in the animal mode of action plausible in humans?" is "no." Since the conclusion is that there is no potential cancer hazard to humans, it can also be concluded that there is no renal cancer risk.

Atrazine produces mammary tumors in female Sprague-Dawley rats but not in F344 rats. A concordance table on qualitative aspects of the key events for mammary tumors produced by atrazine, with columns for both Sprague-Dawley and F344 female rats, and a column for humans, is instructive (Table 1). The specific information that distinguishes the F344 from the Sprague-Dawley rat can be used in evaluating similar processes in humans. In Sprague-Dawley rats, atrazine affects the hypothalamus, leading to inhibition of the LH surge during the estrous cycle. This results in persistent secretion of estrogen and prolactin, ultimately leading to mammary tumors. These hormonal changes do not occur in F344 female rats or in CD-1 mice, a species that also is resistant to the mammary carcinogenic activity of atrazine. Even if the human hypothal-

amus were affected by atrazine in a manner similar to the Sprague-Dawley rat, a totally different result, mainly, a hypoestrogenic state, would be expected. Thus, it can be concluded that the MOA for atrazine-induced mammary tumors in Sprague-Dawley rats does not apply to humans, and, like *d*-limonene, there is neither cancer hazard nor risk to humans for this tumor.

Phenobarbital-induced *thyroid* tumors in rats pose a different set of circumstances. (Note: To fully evaluate phenobarbital would have required separately addressing the mode of action of liver tumors.) Phenobarbital induces liver enzymes involved in the metabolism of thyroid hormones, leading to decreased levels in the blood. Reduced thyroid hormone results in stimulation of the pituitary to produce more TSH, which induces the thyroid to attempt to produce thyroid hormone to replace the amount being metabolized and to enhance follicular cell proliferation, eventually leading to follicular cell hyperplasia, adenomas, and carcinomas. This same liver metabolizing process and pituitary-thyroid axis is present in humans, so on a qualitative basis, there is the potential that this mode of action could be operative in humans. However, because of differences in the pharmacokinetics of thyroid hormones between rats and humans, doses of phenobarbital in humans, although inducing the enzymes leading to the metabolism of thyroid hormones, do not result in a sufficiently decreased level of thyroid hormones to produce a TSH increase. Neither phenobarbital nor other liver enzyme inducers have been shown to increase TSH or to lead to changes in follicular cells in humans. Also, increases in TSH in humans due to other mechanisms frequently lead to an increase in thyroid follicular colloid formation and follicular hypertrophy rather than follicular cell proliferation and hyperplasia. Thus, on both quantitative kinetic and dynamic bases, the response in rats does not occur in humans, even though on a qualitative basis the human has the potential. One can thus conclude on a quantitative basis that this MOA does not apply to humans, and like *d*-limonene and atrazine, neither a carcinogenic hazard nor risk to humans is likely.

The remaining three chemicals, ethylene oxide, chloroform and melamine, all have modes of action that are qualitatively and quantitatively possible in humans, thus indicating the relevance of their modes of action to humans. However, although all three chemicals result in the same answers to the three questions we are posing, the implications for risk assessment are quite different.

For ethylene oxide, the MOA is believed to be based on DNA reactivity, the formation of DNA adducts, leading to DNA mutation and an increased risk of cancer. Obviously, this is possible in both animals and in humans, both qualitatively and quantitatively. Although there may be differences in the quantitative levels in animals versus humans, nevertheless, the process is possible in humans. Thus, a full risk assessment taking this into account in humans is essential. Differences in toxicokinetics between animals and humans can be best taken

into account utilizing validated physiologically based pharmacokinetic (PBPK) models.

Chloroform produces kidney and liver tumors in rodents, and the MOA is similar in both tissues, involving oxidative metabolism of chloroform, leading to cytotoxicity and regenerative proliferation. Regenerative proliferation goes on to liver or kidney tumor formation in rodents. This sequence of events of metabolism and cytotoxicity occurs both in rodents and in humans. However, the sequence of events involves a process producing cytotoxicity that is considered to reflect a nonlinear response. Without cytotoxicity, tumors are not expected to develop. Taking into account differences in the kinetics with a PBPK model, one can reasonably project the levels required to produce toxicity in humans and to estimate levels that might be required to produce tumors. Thus, the approach to risk assessment for chloroform would vary from that for a DNA-reactive carcinogen such as ethylene oxide.

Melamine offers still a different kind of example. Melamine administered at sufficiently high doses produces lower urinary tract calculi in rats, which produce urothelial damage, consequent regeneration, and bladder tumor formation. This is clearly a threshold phenomenon, requiring a dose of melamine sufficient to produce calculi. A similar sequence of events is possible in humans, although the response to calculi is quantitatively significantly less in humans with respect to carcinogenic risk, compared to rodents, especially rats. While little is known regarding the actual chemical exposure in humans, there is considerable information available regarding the formation of calculi in the human urinary tract and the resulting response of the urothelium. Based on both qualitative and quantitative assessments, this MOA is possible in humans. Although melamine can be considered to pose a potential cancer hazard to humans in principle, actual risk is dependent entirely on level of exposure. It has been estimated that the exposure to melamine at currently estimated levels may be 4–5 orders of magnitude lower than that which would be required to produce calculi. If that is the case, melamine does not pose a carcinogenic risk to humans. This distinction is important, and clarity in the distinction of mode of action in hazard characterization, separate from exposure analysis and risk assessment, is a critically important contribution of the framework to increasing transparency and common understanding.

During the course of this project, a second ILSI RSI working group examined current data and analyses on the modes of action of a subclass of peroxisome proliferating chemicals, with a focus on characterizing the animal MOAs for peroxisome proliferator-activated receptor α (PPAR α)-agonists in liver, pancreatic acinar cell, and Leydig cell tumors. As reported in a companion paper (Klaunig *et al.*, 2003), that group successfully applied the new human relevance framework to chemicals operating through the PPAR α mode of action. For pancreatic acinar cell and testicular Leydig cell tumors in rats, the MOA was considered to not be sufficiently delineated yet, so the answer to question 1 is “no”. In contrast, the MOA for

rodent liver tumors was considered to be adequately defined. Based on a variety of investigations, it was also concluded that, based strictly on qualitative considerations, the MOA was relevant to humans. However, based on quantitative toxicodynamic data, it was concluded that the answer to question 3 is "no," that it is not relevant to humans, similar to the conclusion regarding phenobarbital-induced thyroid tumors.

SUMMARY

Current EPA and IPCS guidance offers excellent starting points for animal tumor MOA analysis. To complement and supplement this existing guidance, additional elements of a framework analysis have been developed, on the basis of case studies, to transparently evaluate the relevance of the animal MOA to humans. This led primarily to the establishment of concordance analyses between the species producing tumors and humans. The expanded framework takes into account both qualitative and quantitative aspects of the MOA, addresses the issue of the kinds of data that can be used in evaluating the MOA in humans, and provides a disciplined, transparent approach to comparing the key events of the MOA in laboratory animals to humans. It also identifies critical data gaps in assessing relevance to humans of postulated modes of action.

Although not addressed specifically by the present working group, it became apparent that other end points of toxicity (e.g., manifestations of reproductive, developmental, or neurotoxicity) might be addressed by this framework in similar fashion, because several of the modes of action examined in the first exercise involved intermediate, noncancer events. Another working group has been formed to address this hypothesis.

ACKNOWLEDGMENTS

This ILSI Risk Science Institute project was supported by funding from several offices of the U.S. Environmental Protection Agency and the Existing Substances Division of Health Canada.

Members of this committee came from academia, private industry, and government, and their affiliations are clearly stated on the title page. In addition, those from academia are consultants for a variety of companies as well as serving on several national and international government-related committees and panels.

REFERENCES

- Cohen, S. M., Meek, M. E., Klaunig, J. E., Patton, D. E., and Fenner-Crisp, P. (2003). The human relevance of information on carcinogenic mode of action: Overview. *Crit. Rev. Toxicology* **33**, 581–589.
- Hill, A.B. (1965). The environment and disease: Association or causation? *Proc. R. Soc. Med.* **58**, 295–300.
- Klaunig, J. Babich, M. A., Baetcke, K. P., Cook, J. C., Corton, J. C., David, R. M., DeLuca, J. G., Lai, D. Y., McKee, R. H., Peters, J. M., *et al.* (2003). PPAR α agonist-induced rodent tumors: Mode(s) of action and human relevance. *Crit. Rev. Toxicol.* **33**, 655–780.
- Meek, M. E., Bucher, J. R., Cohen, S. M., Dellarco, V., Hill, R. N., Lehman-McKeeman, L. D., Longfellow, D. G., Pastoor, T., Seed, J., and Patton, D. E. (2003). A framework for human relevance analysis of information on carcinogenic modes of action. *Crit. Rev. Toxicol.* **33**, 591–653.
- Sonich-Mullin, C., Fielder, R., Wiltse, J., Baetcke, K., Dempsey, J., Fenner-Crisp, P., Grant, D., Hartley, M., Knaap, A., Kroese, D., *et al.* (2001). IPCS conceptual framework for evaluating a MOA for chemical carcinogenesis. *Regul. Toxicol. Pharmacol.* **34**, 146–152.
- U.S. EPA (1999). Guidelines for carcinogen risk assessment. Risk Assessment Forum. SAB review draft. U.S. Environmental Protection Agency, Washington, DC. (www.epa.gov/ncea/raf/crasab.htm).