Regulatory Toxicology and Pharmacology 73 (2015) 829–833

Contents lists available at ScienceDirect

Regulatory Toxicology and Pharmacology

journal homepage: www.elsevier.com/locate/yrtph

Commentary

Contrasting directions and directives on hazard identification for formaldehyde carcinogenicity

Lorenz R. Rhomberg

Gradient, 20 University Road, Cambridge, MA 02138, USA

ARTICLE INFO

Article history:
Received 14 October 2015
Accepted 15 October 2015
Available online 19 October 2015

Keywords:
Formaldehyde
Carcinogenicity
National Research Council
Risk assessment
Report on Carcinogens

1. Introduction

In response to a Congressional mandate, the National Research Council (NRC) appointed a committee to review a draft carcinogenicity assessment for formaldehyde developed by the Integrated Risk Information System (IRIS) of the US Environmental Protection Agency (EPA; US EPA, 2010). In their review (NRC, 2011), this committee outlined a "roadmap for reform" for improving methods for critically assessing the scientific literature, intended to apply not just to the formaldehyde assessment but to all future IRIS assessment documents. Subsequently, EPA has been working on reforming its assessment methods in accordance with these recommendations. A revised IRIS report for formaldehyde has not yet been released. Also as mandated by Congress, NRC more recently appointed a separate committee to review the US National Toxicology Program (NTP) entry on formaldehyde in its 12th Report on Carcinogens (RoC; NTP, 2011a). This committee critically examined the RoC listing of formaldehyde as carcinogenic, including NTP's underlying analyses, and further undertook its own evaluation of the evidence while applying its own elaboration of NTP's evaluation and listing criteria (NRC, 2014a).

While each of these two NRC committees responded to somewhat different charges, both examined the same central topic: whether formaldehyde should be identified as a human carcinogen – and if so, for which specific cancer(s) – based on critical evaluation and interpretation of the scientific literature. Examining these parallel reviews reveals key differences in the approaches, scientific methods, and ultimate criteria used by two different US government authorities in identifying and classifying human carcinogens by specific cancer site. As applied in evaluating the scientific literature on formaldehyde exposure and lymphohematopoietic (LHP) cancers, these differences highlight the potential for arriving at divergent conclusions that stem directly from applying different rules to interpret and integrate available evidence. Harmonization of the two processes would seem warranted, including a re-examination of the philosophical perspectives underpinning each.

The two NRC committees each provided evaluations and advice on how to assemble and interpret scientific evidence, and they both took stances on the standards of evidence that a responsible regulatory agency should consider to conclude that human carcinogenicity has been scientifically established as "known." In view of this commonality, it is striking that in applying their respective approaches, the two NRC reviews (NRC, 2011; NRC, 2014a) differed regarding the bearing of many of the available studies pertinent to judging carcinogenic hazard – and more fundamentally, on the very relevance of different kinds of evidence to cancer hazard classification. The chief point of divergence is that NTP's RoC assessment process (and the development of that process as urged by the NRC committee reviewing the formaldehyde RoC listing (NRC, 2011)) stresses assessing the *strength* of what are taken to be the key human studies; it takes the stance that, if sufficiently strong human studies find carcinogenic effects, the existence of these studies alone justifies a conclusion of causality, without needing support from bioassay or mechanistic lines of evidence, or considering any volume of quality studies that fail to demonstrate such an association. In contrast, EPA's IRIS assessment process stresses evaluating and *integrating* all lines of evidence, including an evaluation of whether and how discordant findings (among studies of similar types as well as across epidemiology, bioassays, and mechanistic studies) affect confidence in a causal conclusion. These divergences raise the prospect that risk managers and the public may be confronted with differing statements about whether and with what certainty science has identified a human cancer

E-mail address: lrhomberg@gradientcorp.com.

http://dx.doi.org/10.1016/j.yrtph.2015.10.012
0273-2300/

hazard.

In this commentary, my purpose is to explore and comment upon this divergence of approaches to assembling evidence for carcinogen hazard identification and for justifying judgments as to when carcinogenic effects from an agent can be considered to be "known." I use the case of formaldehyde as a potential cause of human LHP cancers as an example, both because of the importance of this evaluation in its own right and because available evidence in this case directly illustrates the contrast between the two approaches in how evidence is brought to bear. I have not attempted to provide my own critical review of the evidence and the supportable conclusions (though I have done so elsewhere (Rhomberg et al., 2011)), nor do I critique the reviewed NTP and EPA assessments or their NRC reviews in detail. Rather, my aim is to comment on the consequences of two separate NRC committees urging two different agencies to develop their methods in divergent directions, and further to propose the key areas that must be addressed to harmonize cancer hazard evaluations.

2. The NRC's recommendation of a "strength of evidence" approach for NTP and the RoC

Quoting the RoC listing criteria, NRC's review (2014a) states, "a substance can be listed in the RoC as 'known to be a human carcinogen' if 'there is sufficient evidence of carcinogenicity from studies in humans, which indicates a causal relationship between exposure to the agent, substance, or mixture, and human cancer.'" NRC (2014a) further notes, "evidence in experimental animals and a known mechanism of action can provide supporting evidence, but that information is not required by the RoC listing criteria in making a listing recommendation that a substance is known to be a human carcinogen" (NRC, 2014a p. 6). The review committee found that what should constitute "sufficient" human evidence was not well defined. Therefore, they undertook to expand beyond what the RoC criteria have to say on the subject. They classified the available human studies as "strong, moderately strong, or weak" based on their judgments about experimental design, exposure characterization, and the scope for influence of bias or confounding. They then defined "sufficient evidence" in humans as "consistent evidence from two or more strong or moderately strong studies with varied study designs and populations that found an association between exposure to formaldehyde and a specific cancer type for which chance, bias, and confounding factors could be ruled out with reasonable confidence" (NRC, 2014a). This statement itself is not entirely clear as to whether this consistency of evidence must be demonstrated only across the two or more selected studies or across the entire body of epidemiologic evidence that has achieved a "strong" or "moderately strong" rating. That is, is it enough that a set of strong or moderately strong studies agree that incidence of a specific cancer type is increased in people exposed to formaldehyde, even if other studies that achieve similar quality ratings fail to demonstrate the association? In practice, however, the committee identified studies that achieved such quality ratings but demonstrated no association between formaldehyde exposure and LHP cancers in humans, and they nonetheless concluded that the evidence for LHP cancers in humans was sufficient.

The NRC review further notes that the NRC panel was charged to "integrate the level-of-evidence conclusions, and considering all relevant information in accordance with the RoC listing criteria, make an independent listing recommendation for formaldehyde and provide scientific justification" (NRC, 2014a). In carrying out this charge, the review committee found that "the term integrate does not have a standard definition" and that any such integration is only required across human studies, but "a known mechanism is not required." Thus, the NRC review panel (NRC, 2014a) took the stance that, if two or more sufficiently "strong" human studies show a particular effect, this constitutes a direct and sufficient demonstration of human carcinogenicity, even if other human studies fail to show the effect. Furthermore, in such a case, animal and mechanistic information need not be considered in drawing conclusions regarding human carcinogenicity.

3. The NRC's recommendation of an integration and weight-of-evidence approach for EPA and IRIS

Being specific about EPA's revised analysis of formaldehyde and LHP cancers is hard, because the reassessment is in progress. The agency has held a public workshop on the science, however, (on April 30-May 1 of 2014, in Arlington, VA (US EPA, 2014)), which suggested that its new approach will consider more types of data and will embrace an evidence-integration approach. The workshop included a discussion of several hypotheses about possible modes of action and how they might be evaluated for plausibility. On a more general level, the agency has repeatedly expressed its intent to take up and broadly apply the recommendations of the "roadmap for reform" (as set out by the NRC review of its earlier formaldehyde assessment draft (NRC, 2011)) and has shown a great deal of activity along these lines, reflected in changes in the IRIS process and in evolving risk assessment documents for other chemicals and toxic effects. It welcomed the favorable view of its progress as expressed in the NRC (2014b) review of the reforms thus far. That NRC (2014b) review sets out a well-stated expectation of the nature of evidence integration, saying:

> "Rather than organize the narrative around a checklist of criteria, such as the Hill criteria, EPA might consider organizing the narrative as an argument for or against hazard on the basis of available evidence. It should be qualified by explicitly considering alternative hypotheses, uncertainty, and gaps in knowledge… [I]t might begin by considering the conclusions supported by the human evidence and then consider how the available animal evidence confirms, does not support, or is irrelevant to the conclusions. Mechanistic evidence, if available, should be used in the discussion of the animal evidence to determine whether the animal evidence is relevant to the claim about human hazard. Gaps in knowledge and important uncertainties should be explicitly included …. Where the narratives are particularly effective, they explain specifically how different strands of evidence connect (NRC, 2014b)."

4. Assessments of formaldehyde carcinogenicity

The history of official assessments (and their publicly discussed drafts) by various bodies, including the International Agency for Research on Cancer (IARC), EPA through its IRIS program, and NTP through its RoC, shows that to varying degrees, each has shifted its stance over time on the question of formaldehyde carcinogenicity, and, in particular, on the ability of formaldehyde to cause LHP cancers. In part, this shift reflects the advent of new studies and the evolving scientific interpretation of reported findings. On the one hand, some epidemiologic studies associate elevations of certain hematopoietic cancers (at least for groupings of LHP malignancies such as "all leukemias" or "myeloid leukemias") to occupational exposure, at least for some (but not all) ways that exposure levels have been described and classified (*e.g.*, as "peak" *vs.* cumulative exposure). On the other hand, several well-conducted epidemiologic studies with substantial exposures showed no such effects. The key "positive" human study (Beane Freeman et al., 2009) suggests some exposure-response trend within the study population, yet no