Journal of Law and Policy

Volume 15
Issue 1
SCIENCE FOR JUDGES VII:
Evaluating Evidence of Causation & Forensic
Laboratories: Current Issues and Standards

Article 5

2007

Determining Disease Causality From Experimental Toxicology Studies

Ronald L. Melnick

John R. Bucher

DETERMINING DISEASE CAUSALITY FROM EXPERIMENTAL TOXICOLOGY STUDIES

Ronald L. Melnick, Ph.D. and
John R. Bucher, Ph.D.[*]

INTRODUCTION

In contrast to attempts by journals and scientific review panels to reveal and/or limit conflicts of interest in scientific publications and peer review evaluations,[1] conflicts of interest are an inherent component of science-based litigation, and generally include presentations and interpretations of studies that are fashioned to appear consistent and favorable to the position of the sponsoring party.[2] This situation puts an enormous burden on judges and juries, forcing them to wade through disguised biases in order to decipher assertions from facts. Numerous conflicts have and continue to arise over the reliability and interpretation of health effects research data because of uncertainties in the precise extrapolation of adverse effects observed in experimental animals to individual human risk and in the costs and benefits associated with reduction or elimination

[*] Ronald L. Melnick is Director of Special Programs, Environmental Toxicology Program (ETP), National Toxicology Program (NTP), National Institute of Environmental Health Sciences (NIEHS); John R. Bucher is Deputy Director ETP, NTP/NIEHS.

[1] *See, e.g.*, Frank Davidoff et al., *Sponsorship, Authorship, and Accountability*, 345 N. ENG. J. MED. 825, 825-27 (2001); Vincent James Cogliano et al., *The Science and Practice of Carcinogen Identification and Evaluation,* 112 ENVTL. HEALTH PERSP. 1269, 1269-74 (2004).

[2] *See, e.g.*, Wendy Wagner, *The Perils of Relying on Interested Parties to Evaluate Scientific Quality*. 95 Am. J. Pub. Health S99, S99-S106 (2005).

of human exposures to toxic agents.[3] Conflicting opinions on the utility of data obtained from studies in laboratory animals seem to originate largely from concerns of impacts of exposures on human health versus impacts on costs and profits.[4]

Although science seeks to expand our knowledge of facts and truths through the principles of hypothesis generation and hypothesis testing, the way in which our knowledge grows and reflects the truth depends on how questions are framed, how rigorously hypotheses are tested, and to what extent assertions extend beyond actual findings and are portrayed as established facts. This paper focuses on principles of design and evaluation of animal carcinogenicity studies and the utility of these studies for determining disease causality and estimating human cancer risk. Part I will discuss the use of animal studies in the context of public health decisions. Part II will examine issues concerning the design of experimental studies. Part III will explore issues relating to how experimental studies are evaluated and Part IV will explore issues concerning the assessment of human cancer risk based on results from animal experimentation.

I. USE OF ANIMAL STUDIES FOR PUBLIC HEALTH DECISIONS

In 1978, the Department of Health Education and Welfare (now the Department of Health and Human Services) created the National Toxicology Program (NTP), an interagency resource to coordinate toxicology testing programs within the federal government.[5] The mission of the NTP is to strengthen the scientific basis for risk assessments, develop and validate improved testing methods, and provide information about potentially toxic chemicals to health, regulatory, and research

---

[3] Ronald L. Melnick, *A Daubert Motion–A Legal Strategy to Exclude Essential Scientific Evidence in Toxic Tort Litigation*, 95 Am. J. Pub. Health S30, S30-S34 (2005).

[4] Laura Heinzerling & Frank Ackerman, *Pricing the Priceless: Cost-Benefit Analysis of Environmental Protection*, GEO. ENVTL. L & POL'Y INST., GEO. U.L. CENTER (2002).

[5] J.A. Califano, Establishment of a National Toxicology Program, Federal Reg. 43, 53080-53081 (1978).

agencies, the scientific and medical communities, and the public.[6] Since the time of its inception, more than 600 agents have been tested for carcinogenic activity in laboratory animals, and technical reports of those studies have undergone peer review.[7] The standardized approaches used by the NTP for the evaluation of carcinogenicity of environmental and occupational agents are frequently referred to as the gold standard for carcinogen identification.[8] Congress also mandates that the NTP provide and update a list of substances that are "reasonably anticipated" or "known" human carcinogens.[9] Currently, the NTP's Eleventh Report on Carcinogens lists 246 agents under one of these two categories.[10]

Why are animal models used to evaluate human risk? The most obvious explanation is that it is unethical to test for adverse health effects, such as cancer, in humans through intentional exposures.[11] Just as animal models are used in preclinical trials of new pharmaceutical agents before testing in humans, experimental studies performed on animals have been used to assess potential health risks of toxic and carcinogenic agents in our workplace and general environment.[12] The predictive value of animal studies is based on species similarities in the biological processes of disease induction.[13] Another major advantage of

---

[6] *Id.*

[7] *See* Department of Health and Human Services, The National Toxicology Program, http://ntp.niehs.nih.gov (listing toxicology studies).

[8] John R. Bucher, *The National Toxicology Program Rodent Bioassay: Designs, Interpretations and Scientific Contributions*, Annals N.Y. Acad. Sci. 982, 198-207 (2002).

[9] Public Health Services Act, 42 U.S.C. § 241(b)(4) (1999).

[10] NAT'L TOXICOLOGY PROGRAM, U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES, REPORT ON CARCINOGENS at I-5 (11th ed. 2004).

[11] C. Oleskey et al., Pesticide *Testing in Humans: Ethics and Public Policy*, 112 ENVTL. HEALTH PERSP. 914, 914-19 (2004).

[12] *See* Environmental Protection Agency, IRIS Database for Risk Assessment, *available at* http://www.epa.gov/iris (containing a database of toxicology studies).

[13] H.R. Pohl et al., *Risk Assessment of Chemicals and Pharmaceuticals in the Pediatric Population: A Workshop Report*, 42 REGULATORY TOXICOLOGY & PHARMACOLOGY 83, 83-95 (2005); I.H. Russo & J. Russo,

animal studies is the elimination of the need to wait for a high incidence of human cancers, which may take as much as 30 years from time of first exposure to clinical manifestation of disease, before implementing public health protective strategies.[14]

Human epidemiology studies typically have limited exposure information especially at times early in tumor development, and confounding factors are not always known.[15] In contrast, exposure conditions can be finely controlled in experimental studies, making it easier to interpret and assign causality.[16] The major disadvantages of animal studies are that they require extrapolations across species and dose, and they do not capture the full range of human susceptibility due to differences in genetics, health status, diet, life style, and exposures to other agents.[17]

Public health agencies, including the International Agency for Research on Cancer (IARC),[18] NTP,[19] and the U.S. Environmental Protection Agency (EPA),[20] have endorsed the perspective that "in the absence of adequate data in humans, it is

---

Mammary Gland Neoplasia in Long-Term Rodent Studies, 104 ENVTL. HEALTH PERSP. 938, 938-67 (1996).

[14] *See* World Health Organization International Agency for Research on Cancer, *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*, at 14, *available at* http://monographs.iarc.fr/ENG/Preamble/CurrentPreamble.pdf (last visited Dec. 14, 2006).

[15] L.T. Stayner & R.J. Smith, *Methodologic Issues in Using Epidemiologic Studies of Occupational Cohorts for Cancer Risk Assessment*, 14 EPIDEMIOL. PREV. 32, 32-39 (1992).

[16] L. Tomatis, *Role of Experimental and Epidemiological Evidence of Carcinogenicity in the Primary Prevention of Cancer*, 42 ANNALI DELL'ISTITUTO SUPERIORE DI SANITÀ 113, 113-17 (2006) (Italy).

[17] R.J. Preston, *Extrapolations are the Achilles Heel of Risk Assessment*, 589 MUTATION RESEARCH 153, 153-57 (2005).

[18] 77 INT'L AGENCY FOR RESEARCH ON CANCER, *Di(2-ethylhexyl)phthalate, in* IARC MONOGRAPHS ON THE EVALUATION OF CARCINOGENIC RISKS TO HUMANS 41-148 (2000).

[19] NAT'L TOXICOLOGY PROGRAM, *supra* note 10.

[20] U.S. ENVTL. PROT. AGENCY, PUBL'N NO. EPA/630/P-03/001F, GUIDELINES FOR CARCINOGEN RISK ASSESSMENT (2005).

biologically plausible and prudent to regard agents and mixtures for which there is sufficient evidence of carcinogenicity in experimental animals as if they presented a carcinogenic risk to humans." This position is based on the fact that all known human carcinogens that have been studied adequately in experimental animals produce positive carcinogenic results.[21] Hence, even in the absence of adequate human data, public health agencies have classified agents as possibly/probably or likely to be carcinogenic to humans[22] or reasonably anticipated to be a human carcinogen[23] if there is sufficient evidence in animals demonstrating either (1) increased incidence of malignant or malignant and benign tumors combined in two or more species or at multiple sites, or (2) increased incidence in two or more independent studies in one species, or (3) increased incidence in a single study in one species if malignant tumors occur to an unusual degree in incidence, site, type, or age of onset.[24] Several agents that were considered to be possible or probable human carcinogens based on animal data were later confirmed as human carcinogens when reliable epidemiology data (usually occupational exposures) became available (e.g., 1,3-butadiene, cadmium, diethylstilbestrol, ethylene oxide, formaldehyde, and vinyl chloride).[25]

---

[21] INT'L AGENCY FOR RESEARCH ON CANCER, *supra* note 18.

[22] INT'L AGENCY FOR RESEARCH ON CANCER, *supra* note 18; U.S. ENVTL. PROT. AGENCY, *supra* note 20.

[23] NAT'L TOXICOLOGY PROGRAM, *supra* note 10.

[24] INT'L AGENCY FOR RESEARCH ON CANCER, *supra* note 18; NATIONAL TOXICOLOGY PROGRAM, *supra* note 10; U.S. ENVTL. PROT. AGENCY, *supra* note 20.

[25] James Huff, *Chemicals and Cancer in Humans: First Evidence in Experimental Animals,* 100 ENVTL. HEALTH PERSP. 201 (1993), *available at* http://www.pubmedcentral.com/picrender.fcgi?artid=1519590&blobtype=pdf (last visited Nov. 29, 2006)

II. EXPERIMENTAL DESIGN ISSUES

The design of experimental studies is critical for the identification of adverse health effects caused by specific environmental agents and the characterization of dose-response relationships.[26] For example, deficiencies in early studies of benzene in animals, which included too few animals, lack of controls, short study duration, and low levels of exposure, failed to detect carcinogenic effects, even though epidemiology studies had demonstrated a causal association between benzene exposure and leukemia in humans. Subsequent studies that were better designed established benzene as a potent, multi-site carcinogen in rats and mice.[27]

Experimental design issues that might influence the outcome of a carcinogenicity study in laboratory animals are discussed below and include the purity and stability of the chemical agent, the animal models used, the number of animals per dose group, and the exposure levels.

A. Chemical

To ensure that the agent under study is responsible for any observed effects and that any contaminants are not the cause or modifier of that response, the chemical should be tested at high purity. Prior to exposing animals to the agent, it is necessary to demonstrate its stability under the conditions of exposure and storage. If the agent degrades or evaporates during exposure, the accuracy of the targeted administered dose is compromised and

---

[26] *See* Bucher, *supra* note 8.

[27] Cesare Maltoni et al., *Benzene, an Experimental Multipotential Carcinogen: Results of the Long-Term Bioassays Performed at the Bologna Institute of Oncology,* 82 ENVTL. HEALTH PERSP. 109 (1989), *available at* http://www.pubmedcentral.com/picrender.fcgi?artid=1568122&blobtype=pdf (last visited Nov. 29, 2006); J.E. Huff et al., *Multiple-Site Carcinogenicity of Benzene in Fischer 344 Rats and B6C3F$_1$ Mice,* 82 ENVTL. HEALTH PERSP. 125 (1989), *available at* http://www.pubmedcentral.com/picrender.fcgi?artid=1568117&blobtype=pdf (last visited Nov. 29, 2006).

degradation products may contribute to any observed response. For example, a study of trichloroethylene mixed in feed would not have reliable exposure data due to the volatile loss of this chemical from the feed samples and the potential cross contamination of the control groups by inhalation exposure.

B. Animal Models

Rats and mice are the two species most frequently used in tests for carcinogenic activity because they have life spans of about two-and-a-half years[28] and studies of up to 1,000 animals can be performed in reasonably sized animal rooms. Strains of animal models used should have good longevity, genetic stability, and few spontaneous diseases that might shorten their life span, mask any chemical induced effects, or impair metabolism/elimination of the test agent.[29] It is difficult to detect a chemically induced response in an organ with a high spontaneous tumor rate.[30] Both sexes of two species are typically used to identify any sex-specific responses and to confirm multiple species effects.[31]

A major shortcoming of the rodent cancer bioassay is its limited statistical power to estimate the true response rate.[32] Power is the probability of detecting an effect (rejecting the null hypothesis[33]) when an effect exists; it is influenced by the

---

[28] H.A. Solleveld et al., *Natural History of Body Weight Gain, Survival, and Neoplasia in the F344 Rat*, 72 J. NATL. CANCER INST., 929-40 (1984).

[29] G.N. Rao et al., *Mouse Strains for Chemical Carcinogenicity Studies: Overview of a Workshop,* 10 FUNDAMENTALS OF APPLIED TOXICOLOGY 385, 385-94 (1988).

[30] Helmut Greim et al., *Evaluation of Historical Control Data in Carcinogenicity Studies*, 22 HUM. EXP. TOXICOL. 541, 541-49 (2003).

[31] *See* Bucher, *supra* note 8.

[32] Joseph K. Haseman, *Statistical Support of the Proposed National Toxicology Program Protocol*, 11 TOXICOLOGICAL PATHOLOGY 77, 77-82 (1983).

[33] In this case, the null hypothesis is that there is no difference in response between exposed and unexposed animals. Rejection of the null hypothesis implies that an alternative hypothesis is more appropriate (i.e., the response is different in exposed versus unexposed animals).

sample size, the background rate, and the magnitude of the true response.[34] For example, in order to achieve statistical significance at the p 0.05 level (i.e., the odds are 1 in 20 or less that a result occurred simply by chance) in a bioassay consisting of about 50 animals per sex and species per dose group, an incidence of 14 percent (7/50) or greater would be necessary in the exposed group if the incidence in the control group was two percent (1/50). Likewise, if the control incidence is 20 percent (10/50), the incidence in the treated group would need to be 38 percent or greater (19/50) to achieve statistical significance at p 0.05. However, if animal group size were only 25, then an incidence of 40 percent (10/25) would not be significant if the control incidence was still 20 percent (5/25). The power limitation of a study may lead to conflicting opinions when no changes or non-significant elevations in incidence are detected in treatment groups of small size. If the null hypothesis is not true (i.e., exposure produces an effect), the power of the study should be sufficiently large to allow rejection of that hypothesis. A negative, underpowered study provides no assurance of the absence of health risks in exposed populations.

C. Exposure

Because of the limited statistical power of the bioassay when group size is only about 50 animals per sex per dose, high doses are typically used to identify potential carcinogenic hazards and multiple dose-groups are used to characterize dose-response relationships.[35] Unless group size is extremely large (i.e., several hundred to thousands of animals per group), a negative carcinogenicity study (no statistically significant exposure-related effects) conducted at environmental exposure levels can lead to misinterpretation of the carcinogenic potential of an agent. For

---

[34] Joseph K. Haseman, *Statistical Issues in the Design, Analysis and Interpretation of Animal Carcinogenicity Studies*, 58 ENVTL. HEALTH PERSP. 385, 385-92 (1984).

[35] Richard A. Griesemer, *Dose Selection for Animal Carcinogenicity Studies: A Practitioner's Perspective*, 5 CHEM. RES. TOXICOL. 737, 737-41 (1992); Bucher, *supra* note 8.

example, if the true response rate at an environmental exposure level is 1 per 1000, then this exposure would not reveal a significant response if the animal group size is 50. However, environmental exposures associated with increased cancer risk of 1 per 1000 in human populations would be dreadful. Data from preliminary studies, typically of 4 to 13 weeks duration, including evaluations of body weight changes, clinical observations, and pathological effects are used to estimate the maximally tolerated dose or the minimally toxic dose (MTD) for the subsequent cancer study.[36]

The identification and use of an MTD is a critical aspect in the design of experimental carcinogenicity studies of low statistical power.[37] The MTD should not cause increases in mortality other than from chemically induced tumors. Lower doses (one-half MTD and one-quarter to one-tenth MTD) are used if the high dose selected for the chronic study is found to be too high (excessive mortality) and to provide dose-response information.[38] Pharmacokinetic information is also used to ensure that no more than one of the selected doses is above a level that saturates the processes of absorption, metabolic activation, or detoxification.[39] In the absence of human data, the dose-response relationship in an experimental animal study serves as the basis for estimating risks at human exposure levels.[40] A much better characterization of the true dose-response can be achieved with several dose groups rather than

---

[36] Victor A. Fung et al., *The Carcinogenesis Bioassay in Perspective: Application in Identifying Human Cancer Hazards*, 103 ENVTL. HEALTH PERSP. 680, 680-83 (1995).

[37] C.J. Portier & D.G. Hoel, *Design of Animal Carcinogenicity Studies for Goodness-of-Fit of Multistage Models*, 4 FUNDAMENTALS OF APPLIED TOXICOLOGY 949, 949-59 (1984).

[38] John R. Bucher et al., *Workshop Overview, National Toxicology Program Studies: Principles of Dose Selection and Applications to Mechanistic Based Risk Assessment,* 31 FUNDAMENTALS OF APPLIED TOXICOLOGY 1, 108 (1996).

[39] J.R. Buchanan et al., *Purpose and Guidelines for Toxicokinetic Studies Within the National Toxicology Program*, 105 ENVTL. HEALTH PERSP. 468, 468-71 (1997).

[40] U.S. ENVTL. PROT. AGENCY, *supra* note 20.

only two widely spaced dose groups.

Typical carcinogenicity studies in rats and mice involve exposures beginning at six weeks of age and continuing for two years; this exposure period corresponds roughly with early adulthood through most of an occupational life span.[41] However, because of data indicating greater susceptibility with exposure to mutagens[42] and endocrine disruptors[43] during growth and early developmental stages, earlier periods of exposure are frequently included in animal carcinogenicity studies when fetal or childhood exposure to such agents might occur. The two-year duration limit was selected to minimize late-developing background tumor responses in controls as well as in animals exposed to the test agent that might preclude the detection of chemical-induced effects.[44] Studies with exposure durations shorter than two years are also problematic because of their reduced sensitivity to detect increases in late-appearing tumors that are related to treatment with the test agent.[45]

III. EVALUATION ISSUES

The conduct and evaluation of a cancer bioassay requires a multidisciplinary effort, including expertise from toxicologists, laboratory animal veterinarians, chemists, histologists, pathologists, cellular/molecular biologists, and statisticians. The interpretation of tumor data may be affected by the thoroughness of the histopathology evaluations, methods of statistical analysis,

---

[41] Joseph K. Haseman et al., *Carcinogenesis Bioassays: Study Duration and Biological Relevance,* 39 FOOD & CHEMICAL TOXICOLOGY 739, 739-44 (2001).

[42] J. M. Rice et al., *Comparative Transplacental Carcinogenesis by Directly Acting and Metabolism-dependent Alkylating Agents in Rodents and Nonhuman Primates*, 96 IARC SCI. PUBL. 17, 17-34 (1989).

[43] A.L. Herbst et al., *Prenatal Diethylstilbestrol Exposure and Human Genital Tract Abnormalities,* 51 NAT'L CANCER INST. MONOGRAPH 25, 25-35 (1979). An endocrine disruptor is a natural or synthetic chemical that may mimic or antagonize the actions of natural hormones responsible for maintaining homeostasis and regulating development.

[44] Solleveld, *supra* note 28.

[45] Haseman, *supra* note 41.

and mechanistic considerations. With respect to human relevance of animal cancer data, the issue of concordance of the site of occurrence of the tumor has been raised; this topic is also discussed below.

A. Histopathology

The detection of toxic effects and neoplastic lesions in animals depends on the thoroughness of the necropsies, or post-mortem examinations of the animal, and the microscopic examinations performed on slides of tissue sections. The NTP requires examination of all organs and tissues, approximately 40 per animal.[46] In some cases, multiple sectioning of an organ in exposed and control groups may be necessary to obtain a truer estimate of the incidence of neoplastic lesions, especially for small lesions that may not be detected at necropsy.[47] Although diagnostic criteria have been established for most observable lesions, it is not unusual for pathologists to disagree in their judgment of lesions, especially those that are part of a continuum of progressive change.[48] After diagnoses by the study pathologist, all NTP studies undergo an independent quality assessment (QA) pathology review.[49] This is followed by a pathology working group review (typically eight to ten pathologists) that seeks to resolve discrepancies in diagnoses between the original and the QA pathologists.[50] Studies that lack multiple pathology evaluations may yield diagnostic data that would not be generally accepted.

---

[46] Department of Health and Human Services, National Toxicology Program, http://ntp.niehs.nih.gov (last visited Dec. 15, 2006).

[47] S.L. Eustis et al., *The Utility of Multiple-section Sampling in the Histopathological Evaluation of the Kidney for Carcinogenicity Studies*, 22 TOXICOLOGIC PATHOLOGY 457, 457-72 (1994).

[48] S.L. Eustis, *The Sequential Development of Cancer: A Morphological Perspective*, 49 TOXICOLOGY LETTERS 267, 267-81 (1989).

[49] G. A. Boorman et al., *Quality Assurance in Pathology for Rodent Carcinogenicity Studies, in* HANDBOOK OF CARCINOGEN TESTING 345, 345-57 (H. Milman & E. Weisburger eds., 1985).

[50] *Id.*

In addition to statistical analyses described below, it should be recognized that other factors may contribute to the interpretation of tumor data, including: (1) the occurrence of uncommon versus common tumors, (2) evidence of progression of lesions, such as benign to malignant where it is appropriate to combine,[51] (3) tumor occurrence with reduced latency, (4) multiplicity in a site-specific tumor response, (5) evidence of metastases, and (6) supporting evidence of proliferative preneoplastic lesions (lesions with high cell replication rates that may progress to tumors) at the same organ site or detection of the same lesion in the other sex or species. Because cancer development is a multi-step process with a long time period between exposure and manifestation of metastatic neoplasia, evidence of enhanced disease progression (e.g., reduced latency, tumor multiplicity, and metastasis) in an animal study is another indication that the agent promotes cancer development.

B. Statistics

If mortality in a dose group is different from that of controls, it is important that pair-wise comparisons and analyses of trends be based on tumor rates that have been adjusted for deaths occurring before the end of the study.[52] The reason for this survival-based adjustment is that if animals died early from causes other than tumors at the organ site of interest, then those animals would not have been on study long enough to provide a full contribution of risk to that study group. Failure to adjust for differences in survival will yield unreliable estimates of cancer risk, and possible misinterpretations of a true site-specific effect.

Although comparisons between the concurrent control group and the exposure groups are the most valid for identifying chemically induced effects, comparisons with historical control

---

[51] E. E. McConnell et al., *Guidelines for Combining Neoplasms for Evaluation of Rodent Carcinogenesis Studies*, 76 J. NAT'L CANCER INST. 283, 283-89 (1986).

[52] Joseph K. Haseman, *Statistical Issues in the Design, Analysis and Interpretation of Animal Carcinogenicity Studies*, 58 ENVTL. HEALTH PERSP. 385, 385-92 (1984).

data may also be helpful in interpreting treatment-related effects.[53] For meaningful comparisons, the conditions of the current study must be similar to those in the historical database and diagnostic criteria must be identical. Thus, comparisons must be specific for the species, sex, and strain of animals, the route of exposure, and the diet. For example, dietary factors may influence animal survival, organ function, and spontaneous or chemically induced tumor rates. Conflicting interpretations of study findings may arise with improper use of historical databases.

C. Mechanistic Considerations

Results from animal or *in vitro* studies that attempt to determine the mode-of-action for disease induction (i.e., mechanistic studies) have been used to upgrade or downgrade cancer risk classifications of agents that have inadequate or limited evidence in humans.[54] For example, based on mechanistic data on early molecular and cellular events that likely contribute to cancer outcome, IARC and NTP upgraded ethylene oxide[55] and 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD)[56] to "known human carcinogens." In both cases, the evidence of carcinogenicity was limited in humans and sufficient

---

[53] Joseph K. Haseman, *Use of Historical Control Data in Carcinogenicity Studies in Rodents*, 12 TOXICOLOGICAL PATHOLOGY 126, 126-35 (1984).

[54] NAT'L TOXICOLOGY PROGRAM, *supra* note 10; 60 INT'L AGENCY FOR RESEARCH ON CANCER, *Ethylene Oxide, in* IARC MONOGRAPHS ON THE EVALUATION OF CARCINOGENIC RISKS TO HUMANS 73-159 (1994); 69 INT'L AGENCY FOR RESEARCH ON CANCER, *2,3,7,8-Tetrachlorodibenzo-para-dioxin, in* IARC MONOGRAPHS ON THE EVALUATION OF CARCINOGENIC RISKS TO HUMANS 33-343 (1997).

[55] NAT'L TOXICOLOGY PROGRAM, *supra* note 10; 60 INT'L AGENCY FOR RESEARCH ON CANCER, *Ethylene Oxide, in* IARC MONOGRAPHS ON THE EVALUATION OF CARCINOGENIC RISKS TO HUMANS 73-159 (1994).

[56] NAT'L TOXICOLOGY PROGRAM, *supra* note 10; 69 INT'L AGENCY FOR RESEARCH ON CANCER, *2,3,7,8-Tetrachlorodibenzo-para-dioxin, in* IARC MONOGRAPHS ON THE EVALUATION OF CARCINOGENIC RISKS TO HUMANS 33-343 (1997).

in experimental animals. The upgrading of ethylene oxide was based largely on the induction of chromosomal aberrations in peripheral lymphocytes, micronuclei in bone marrow cells, and hemoglobin adducts in exposed workers.[57] For TCDD, the upgrading was based on data demonstrating that the multi-site carcinogenicity of this chemical in experimental animals was due to a mechanism involving activation of the aryl hydrocarbon receptor and studies showing that this receptor is highly conserved across species and functions the same way in humans as it does in experimental animals.[58] Thus, even though human epidemiological data alone were not sufficient to reach the "known human carcinogen" classification category, these authoritative bodies considered the combination of animal data and mechanistic data sufficient to support the upgraded "weight-of-evidence" conclusions. The term "weight-of-evidence" is used by the EPA to reflect evaluations that are "based on the combined strength and coherence of inferences appropriately drawn from all of the available information."[59]

Similarly, one should be very concerned of the potential cancer risks for vinyl bromide (VBr) and vinyl fluoride (VF), because they are structural analogs of vinyl chloride, a known human carcinogen. Mechanistic data show that these three chemicals are metabolized to DNA-reactive intermediates by enzymes found in humans[60] producing identical promutagenic DNA adducts. Experimental carcinogenicity studies show that these three chemicals induce multiple tumor types, including

---

[57] NAT'L TOXICOLOGY PROGRAM, *supra* note 10; 60 INT'L AGENCY FOR RESEARCH ON CANCER, *Ethylene Oxide, in* IARC MONOGRAPHS ON THE EVALUATION OF CARCINOGENIC RISKS TO HUMANS 73-159 (1994).

[58] NATIONAL TOXICOLOGY PROGRAM, *supra* note 10; INT'L AGENCY FOR RESEARCH ON CANCER, *2,3,7,8-Tetrachlorodibenzo-para-dioxin, in* 69 IARC MONOGRAPHS ON THE EVALUATION OF CARCINOGENIC RISKS TO HUMANS 33-343 (1997).

[59] U.S. ENVTL. PROT. AGENCY, *supra* note 20, at 2-1.

[60] F. Peter Guengerich et al., *Role of Human Cytochrome P-450 IIE1 in the Oxidation of Many Low Molecular Weight Cancer Suspects*, 4 CHEMICAL RESEARCH IN TOXICOLOGY 168, 168-79 (1991).

angiosarcomas, in rats and mice.[61] Because of these similarities, it would be unethical to require human data on VBr or VF before dealing with these chemicals as known human carcinogens.

On the other hand, IARC downgraded the classification of di(2-ethylhexyl)phthalate (DEHP) from "possibly" to "not classifiable as to its carcinogenicity to humans."[62] The evidence of the carcinogenicity of DEHP was concluded to be sufficient in animals, based on increased incidences of liver tumors in rats and mice, and inadequate in humans.[63] The downgrading of the animal cancer evidence was based on the panel's acceptance of the hypothesis that DEHP induces liver tumors in rats and mice by a non-DNA-reactive mechanism involving an increase in the number of peroxisomes (subcellular structures that contain several oxidase enzymes). The tumor responses in rats and mice were considered to be irrelevant to humans because peroxisome proliferation had not been documented either in human hepatocyte cultures exposed to DEHP or in the liver of exposed non-human primates.[64] However, peroxisome proliferation alone has not been shown to provide a reasonable mechanistic explanation for the different carcinogenic potencies of peroxisome proliferators in the rat liver.[65]

Consequently, peroxisome proliferation may not be a reliable marker for evaluating human cancer risk. Other mode-of-action hypotheses for rodent liver tumor induction by peroxisome proliferators have not been tested. Thus, the mechanistic basis

---

[61] Ronald L. Melnick, *Carcinogenicity and Mechanistic Insights on the Behavior of Epoxides and Epoxide-forming Chemicals*, 982 ANNALS N.Y. ACADEMY OF SCIENCES 177, 177-89 (2002).

[62] INT'L AGENCY FOR RESEARCH ON CANCER, *supra* note 18.

[63] *Id*.

[64] *Id*.

[65] D. S. Marsman et al., *Relationship of Hepatic Peroxisome Proliferation and Replicative DNA Synthesis to the Hepatocarcinogenicity of the Peroxisome Proliferators Di(2-ethylhexyl)phthalate and [4-chloro-6-(2,3-xylidino)-2-pyrimidinylthio]acetic acid (Wy-14,643) in Rats*, 48 CANCER RESEARCH 6739, 6739-744 (1988).

for the IARC decision to downgrade DEHP[66] is not supported by experimental evidence. Public health decisions that could lead to unrestricted use and exposure to carcinogenic agents should not rely on untested hypotheses.[67]

D. Site Concordance

Differences in the organ or tissue in which tumors arise in animals and humans are one of the most contentious issues in the evaluation of the human relevance of experimental carcinogenicity data.[68] This is largely due to the fact that the etiologies or causes of most cancers are not known and the bases for differences in species susceptibility are not fully understood.[69] There are several examples in which human cancer sites correspond with one animal species, but not the second species (e.g., hematopoietic cancers induced by benzene in mice and humans, but not in rats) and there are also examples of site correspondence among rats, mice, and humans.[70]

Although site correspondence may strengthen the animal-to-human association, there are numerous reasons why that correspondence should not be a requirement for human relevance and causality. First, exposure factors might contribute to differences in sites of tumor induction, including the route, frequency, duration, intensity, as well as age at onset of

---

[66] INT'L AGENCY FOR RESEARCH ON CANCER, *supra* note 18.

[67] Lornezo Tomatis, *The IARC Monographs Program: Changing Attitudes Towards Public Health*, 8 INT'L J. OCCUPATIONAL ENVTL. HEALTH 144, 144-52 (2002).

[68] Robert R. Maronpot et al., *Relevance of Animal Carcinogenesis Findings to Human Cancer Predictions and Prevention*, Toxicologic Pathology 32 Suppl. 1, 40-48 (2004).

[69] T.L. Goodrow, *One Decade of Comparative Molecular Carcinogenesis*, Progress in Clinical & Biological Res. 395, 57-80 (1996); Amanda Ewart-Toland & Allan Balmain, *The Genetics of Cancer Susceptibility: From Mouse to Man*, Toxicologic Pathology 32 Suppl. 1, 26-30 (2004).

[70] Ronald L. Melnick & J.E. Huff, *1,3-Butadiene Induces Cancer in Experimental Animals at All Concentrations from 6.25 to 8000 Parts Per Million*, IARC Sci. Publ. 127, 309-322 (1993).

exposure.[71] Exposure conditions are known to affect sites and dose-response for tumor induction in experimental studies. If gestational or early childhood exposure is important for tumor induction, then studies of male workers would not be expected to produce the same response as animal studies that include exposures during these stages of development.

Second, there may not be an adequate human epidemiological study on the agent or an epidemiological study may not have detected a true increase in risk of certain cancers in exposed human populations due to inadequate exposure information, misclassifications, insufficient follow-up, and/or inadequate study power.[72]

Third, human susceptibility to environmental carcinogens varies for a large number of reasons including genetic factors, health status, diet, lifestyle (e.g., smoking, alcohol consumption), age, and other exposures (e.g., medications, occupational experiences).[73] Because cancer development is the likely result of interactions among environmental factors and individual susceptibility factors, differences in sites of tumor induction among individuals are not unusual for known human carcinogens. For example, although the lung is the most common cancer site among cigarette smokers, many people do not develop lung tumors but instead develop cancers of the bladder, kidney, nasal cavity, lip, esophagus, or pancreas. In some smokers no cancer is evident. Numerous host susceptibility factors could account for lack of site correspondence among individuals in exposed populations or between experimental animals and humans.

Because of the much greater genetic diversity in humans compared to strains of laboratory animals, the range of expected

---

[71] U.S. ENVTL. PROT. AGENCY, Publ. No. EPA/630/R-03/003F, Supplemental Guidance for Assessing Susceptibility from Early-Life Exposure to Carcinogens, 24 (2005).

[72] Stayner and Smith, *supra* note 15.

[73] F.P. Perera, *Molecular Epidemiology of Environmental Carcinogenesis*, 154 Recent Results in Cancer Research, 39-46 (1998); R.G. Dumitrescu & I. Cotarla, *Understanding Breast Cancer Risk—Where Do We Stand in 2005?*, 9 J. CELLULAR & MOLECULAR MED. 208, 208-21 (2005).

human response may include subgroups that are less, equal, or more sensitive than the animal models used in the experimental cancer studies. Unless a qualitative difference between animals and humans is clearly shown to be the determinant of a species-specific cancer response, it is prudent to assume that a carcinogenic effect in animals is a reliable indicator of potential cancer risk in humans. This perspective has been endorsed by major national and international public health agencies that evaluate human cancer risks associated with exposures to environmental and occupational agents.[74]

IV. ESTIMATING HUMAN CANCER RISK FROM ANIMAL DATA

Risk assessment provides a systematic approach for characterizing the nature and probability of adverse effects (i.e., health risks) occurring in individuals or populations exposed to hazardous agents and often serves as the basis for risk management decisions on whether and to what extent human exposure to such agents should be controlled. The National Academy of Sciences/National Research Council (NAS/NRC) developed guidelines for the conduct of risk assessments in the federal government.[75] The risk assessment paradigm developed by the NAS/NRC consists of four parts: hazard identification, dose-response assessment, exposure assessment, and risk characterization.[76]

When adequate human data is not available, dose-response data from studies in laboratory animals serve as the basis for estimating risks in exposed humans.[77] A quantitative risk assessment requires conversion of animal doses to human

---

[74] NAT'L TOXICOLOGY PROGRAM, *supra* note 10; INT'L AGENCY FOR RESEARCH ON CANCER, *supra* note 18; U.S. ENVTL. PROT. AGENCY, *supra* note 20.

[75] COMM. ON THE INSTITUTIONAL MEANS FOR ASSESSMENT OF RISK TO PUB. HEALTH, NAT'L RESEARCH COUNCIL, *Risk Assessment in the Federal Government: Managing the Process* 19 (National Academy Press 1983).

[76] *Id*.

[77] U.S. ENVTL. PROT. AGENCY, *supra* note 20.

equivalent doses.[78] If a verified physiologically based pharmacokinetic model is available for the specific agent, this might be used to describe the internalized dose in humans by replacing animal physiological and biochemical parameters in the model with those specific for humans (e.g., breathing volumes, organ sizes, cardiac output, metabolic rate constants, etc.). Computer-based dosimetry models consist of a series of mathematical equations that represent, in quantitative terms, the complex biological processes that affect the absorption, distribution, metabolism, and elimination of the agent in the intact animal. These models can accommodate parameter values that represent the range and distribution of activities that exist in human populations. Reliable models can provide estimates of tissue dose as a function of duration and frequency of exposure to various environmental exposure levels.[79] However, one must be cautious of assumptions in models that have not been validated because they could lead to inaccurate estimates of tissue dose. If a verified dosimetry model is not available, human equivalent exposures are typically obtained by scaling animal doses to humans as a function of body surface area (i.e., body weight to the ¾ power).[80] The assumption in this approach that metabolic activities differ among species according to body surface area may not be true.

The next step in the risk assessment process is to fit empirical models (or mechanistic-based models if the mechanism of the disease induction is known) to the dose and cancer incidence data corrected for background and then to determine which model provides the best fit to the dose-response data.

By extending the dose-response curve, it is possible to estimate the cancer risk at human exposure levels, or to estimate the exposure levels that are associated with specific risks (e.g., one per hundred, one per thousand, one per million). In their most recent cancer risk assessment guidelines, the EPA

---

[78] *Id.*

[79] Michael C. Kohn, *Achieving Credibility in Risk Assessment Models*, 79 TOXICOLOGY LETTERS 107, 107-14 (1995).

[80] U.S. ENVTL. PROT. AGENCY, *supra* note 20.

recommends modeling dose-response data to determine the effective dose associated with one percent (ED01) or 10 percent (ED10) risk and the lower 95 percent limit on those dose estimates (LED01 or LED10).[81] From this point a straight line is drawn to zero risk for mutagenic carcinogens or chemicals for which the mode-of-action has not been characterized. Cancer risks at lower exposure levels are estimated from the slope of this line. These risks can then be compared to age, sex, and race-dependent incidences of specific cancers in the U.S. population reported in the Surveillance, Epidemiology, and End Results (SEER) data set[82] to determine the relative risk of developing or dying from cancer due to a particular exposure. Relative risk is the incidence of disease in an exposed population divided by the incidence of that disease in the unexposed population.[83]

For example, consider women working at a facility where they are exposed until age 60 to a carcinogen at a level associated with additional leukemia risk of 1.5 per thousand. Since the probability of women in the general U.S. population dying from leukemia by age 60 is 1.25 per thousand, then the relative risk of dying from leukemia due to that exposure is 2.2. The estimated incidence in the unexposed population is 1.25/1000, while the estimated incidence in the exposed population is the background risk (1.25/1000) plus the additional risk associated with exposure (1.5/1000). Relative risk equals $(0.00125 + 0.0015)/0.00125 = 2.2$.

---

[81] *Id.*

[82] Surveillance Epidemiology and End Results, National Cancer Institute, http://seer.cancer.gov/studies/endresults/study26.html (last visited Dec. 5, 2006).

[83] Ronald L. Melnick, *Occupational Chemical Carcinogenesis*, in 1 PATTY'S INDUSTRIAL HYGIENE AND TOXICOLOGY 117, 117-67 (E. Bingham, B. Cohrssen, C.H Powell, ed., 2001).

CONCLUSION

Data from properly designed and evaluated studies in experimental animals have been and continue to be reliable sources of information for the identification of potential human health hazards and the estimation of risks in exposed populations. Properly designed animal studies must include: (1) animal models that are sensitive to the endpoints under investigation, (2) detailed characterization of the agent and the administered doses, (3) challenging doses (MTD) and durations of exposure (at least two years for rats and mice) that would allow a reliable determination of whether or not the agent poses a health hazard, (4) sufficient numbers of animals per dose group to have adequate statistical power to detect a true effect, (5) multiple dose groups to allow characterization of dose-response relationships, (6) complete and peer-reviewed histopathological evaluations, and (7) pair-wise comparisons and analyses of trends based on survival-adjusted tumor rates. Mechanistic information and pharmacokinetic models that have been adequately tested may impact the characterization of dose-response relationships. Until the etiology of environmentally induced cancers and the basis for human susceptibility are better understood, site correspondence should not be a requirement for judging human relevance and causality.