Annu. Rev. Pharmacol. Toxicol. 1991. 31:621–52

SCIENTIFIC CONCEPTS, VALUE, AND SIGNIFICANCE OF CHEMICAL CARCINOGENESIS STUDIES*

James Huff, Joseph Haseman, and David Rall

National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709

KEY WORDS: long-term studies, pesticides, human risks, cancer, synthetic and natural chemicals

INTRODUCTION

Principle 1. Effects in animals, properly qualified, are applicable to man (1).

Synthetic chemicals and those naturally occurring chemicals that have become 'man-made' via chemical synthesis and high volume production have only relatively recently been introduced into our way of life, and now have become essential for our adopted and evolved lifestyle. With the growing avalanche of chemicals comes the necessity for more public health prudence to minimize human exposures from these chemicals (derivatives, metabolites, and reaction products) and their myriad formulations during synthesis, manufacture, mixing, packaging, transport, use, and disposal. Toxicologic characterization of chemicals, mixtures of chemicals, and consumer products is essential for deciding policy to best protect public health while retaining the benefits we have come to expect. The profile of biological activities and toxicological effects of these agents or substances can be accomplished by using an array of in vitro and in vivo assays. These experiments are conducted

*The US Government has the right to retain a nonexclusive, royalty-free license in and to any copyright covering this paper.

to help determine the toxicologic patterns and to identify those chemicals considered most likely to represent a potential hazard or risk to humans and to assess or predict such hazards or risks.

The central theme of this paper deals with the scientific value of long-term chemical carcinogenesis experiments using laboratory rodents as reliable predictive surrogates for humans. We therefore evaluated long-term carcinogenesis studies on nearly 400 chemicals, and give our interpretations in appropriate detail to allow readers to form their own opinions. These studies were designed, conducted, and reported by the National Cancer Institute (NCI) into early 1980 and by the National Toxicology Program (NTP) since then.

THE PROBLEM

There are no harmless substances; there are only harmless ways of using substances (2).

In early 1990 the ten millionth unique chemical was synthesized and entered into the Chemical Abstracts Service's computer-based registry. Clearly and thankfully we are not concerned with the vast majority of these laboratory curiosities. Nonetheless a monumental problem exists for determining the toxicological properties on the near 100,000 chemicals that were being used until recently or are being used in large amounts today; add to this structural menagerie the collection of 500–1000 chemicals that each year are newly and intentionally placed into our environmental milieu. Other sources of creative exposures come from chemical changes that take place in our combustion engines and furnaces, in our industrial smokestacks, in our municipal and industrial disposal systems, in our water and sewage treatment plants, and in our modern agricultural practices and home kitchens.

As examples of the sheer magnitude and universality of the problem, imagine the logistics of mining and transporting raw sources of chemicals (e.g. crude oil, ores, intermediates); producing, extracting, synthesizing, and preparing end-system chemicals; transporting, bottling, canning, fabricating, extruding, mixing, and formulating industrial and consumer products; and the huge task of disposing, recycling, removing, or cleaning-up all the resultant solids, liquids, gases, vapors, aerosols, and particulates generated as waste products. At each stage potential chemical hazards of one sort or another should be recognized and appropriate health and safety measures instituted; in many cases such measures have already been introduced, but for a large part have dealt with potential physical injuries and not chemical exposures.

Chemical Production

The upsurge of organic chemical production in the United States is a modern phenomenon: in 1940 about 1,300,000,000 lbs of synthetic chemicals were produced; in 1950 this increased 38-fold to 49 billion lbs; in 1960 the amount produced doubled to 97 billion; in 1970 production more than doubled to 233 billion; in 1980 the volume hit 320 billion; and in 1988 the combined production of all synthetic organic chemicals and primary products from petroleum and natural gas amounted to 400 billion lbs, with sales of 217 billion valued at close to $100,000,000,000 (3). Thus, in less than 50 years the production of organic chemicals has jumped more than 300 times, and continues to increase. Unfortunately, only a proportionately few of these natural, synthetic, or extracted 'natural' chemicals have been adequately studied and evaluated for their toxicologic potential (4).

Pesticides

As an important subclass of end-use chemicals, synthetic pesticides include fungicides, herbicides, insecticides, rodenticides, and related products such as plant growth regulators, seed disinfectants, soil conditioners, soil fumigants, and synergists. The following data are given in terms of 100% active ingredients, and they exclude such materials as diluents, emulsifiers, and wetting agents (5). For this eclectic collection 1.2 bill. lbs were produced, sold (for $4.4 bill.), and used in 1988; this represents about 4.7 lbs of pesticide for each person in the United States. As an estimate of production in 1950 and using the population at that time, each person would have had about one-half as much pesticide. Ten million individuals—including family members and hired workers—in the United States alone are engaged in farming (6). Thus, if the amounts of pesticide produced in 1988 were distributed equally among this farm population each person would have about 120 lbs of agrichemical pesticides.

The three major subcategories and the amounts produced are herbicides (702 million lbs), fungicides (110 million lbs), and insecticides (353 million lbs). There appears to be a relatively consistent and stable trend in pesticide production during the 1980s, each year hovering at about one billion lbs (3).

Carcinogenesis data and interpretations on the agricultural chemicals are reported elsewhere by Huff & Haseman (7), who concluded that recent epidemiology studies have confirmed the link between pesticide exposure and cancer, and that laboratory studies in animals have also identified other pesticides with carcinogenic potential to humans. Thus, due caution must be exercised in the manufacture and use of these chemicals, and in some cases appropriate public health action may be needed to reduce or eliminate human exposure (8).

Natural Versus Synthetic Chemicals

Concern is expressed whether synthetic chemicals or naturally occurring chemicals are most 'guilty' of chemical-associated human exposures and human illness. Additionally, confusion surrounds the definition of naturally occurring or synthetically produced chemicals. This particular issue has been given widespread attention, and has tended to mislead the public by overstating that only natural chemicals in the food we eat are of any harm, whereas synthetically produced chemicals and the chemicals made and used as industrial chemicals present little or no potential hazard to public health. This is not correct, and the rationale put forth that we are exposed mainly to "natural carcinogens" is flawed. As correctly stated by Tomatis et al (9), "the partition between man-made and natural carcinogens [or chemicals] is rather artificial; for instance, tobacco is a natural plant, but cigarettes and tobacco smoke can hardly be called natural products. Another example is asbestos, a naturally occurring mineral, but it is only through mining, milling, factory production and the handling of asbestos products that it is disseminated into the environment, leading to direct human exposure."

To this category of chemicals stated to be naturally occurring but made widely available only through 'synthetic means' can be added the high volume petroleum-based benzene and 1,3-butadiene; the plant-extracted (and now synthetically made) drugs reserpine and ascorbic acid; the mined and milled lead and other metals; and the botanical pesticides nicotine and (now synthetic) pyrethroids and rotenone. In our opinion, both classes of 'chemicals' are without doubt potentially hazardous to public health. With certain exceptions, we do not believe that the normal levels of natural or synthetic chemicals that people typically eat (along with anticarcinogens and antioxidants and many competing chemicals) present any serious or imminent hazard to humans. If an individual or a group of persons confined their eating day-in, day-out to one food that contained potentially harmful natural or synthetic chemicals, or to a food source containing chemicals that might exacerbate a potentially hazardous condition, some risk to human health might accrue.

These illustrations present the issue of naturally occurring versus synthetically produced chemicals, but a complementary dispute regarding human capability to equally detoxify both classes of chemicals needs clarification. Using evolutionary theory, over long stretches of time humans would be expected slowly and surely to adapt to handle metabolically the 'new' chemicals eaten when new foods were discovered (a major adaptation being the perhaps drastic changes introduced with fire for cooking, especially meats). This concept of adaptation has long been considered a logical means of indicating that the avalanche of totally new and diverse chemical structures would present a novel challenge to our biological armamentarium. For example, until humans were blanketed by exposures to the organochlorine insecticides (e.g. DDT) no organic chlorinated (or brominated) hydrocarbons existed in our bodies (we do have two iodinated hydrocarbons); now everyone on earth carries chlorinated hydrocarbons in their adipose tissue.

Table 1 Estimated annual global emissions of selected metals to the atmosphere, circa 1980 (11)

Metal	Human activity	Natural activity	Ratio of human to natural activity
	(000 metric tons)		
Lead	2,000	6	333
Zinc	840	36	23
Copper	260	19	14
Vanadium	210	65	3
Nickel	98	28	4
Chromium	94	58	2
Arsenic	78	21	4
Antimony	38	1	38
Selenium	14	3	5
Cadmium	6	0.3	20

In 1973 Elias (10) explained clearly his view of this phenomenon: "Primitive man learned to cope with a few naturally occurring chemical substances mostly represented by his food and drink, and the ministrations of the medicine man. And he had aeons to refine this process of adaptation. Not so modern man. Exposure to new chemicals in the total environment has escalated on an unprecedented scale and despite the institution of safeguards, and attempts at slowing down this process, it is most unlikely that man will be able to adapt to every new chemical he is likely to meet. Thus toxicology assumes a new and vastly different role in this struggle for survival." Fortunately, man and animals have benefited greatly from the long-term exposures to truly naturally occurring chemicals; the challenges allowed biological systems to develop as early man gradually moved into new environmental milieux. The impacts of human activities on technologic-induced exposures to the naturally occurring metals shows a range of increased global exposures from a high of 333 times for lead to a low of 2 for chromium (Table 1; 11). Similar increases could be shown for a variety of synthetic 'natural' organic chemicals as well.

To re-evaluate our current agricultural practices of crop protection using synthetic pesticides, The American Chemical Society published in their ACS Symposium Series a group of papers on Insecticides of Plant Origin (12). Because many plants have developed over time the capability to defend against marauding insect populations by producing chemicals that are toxic to insects or that even disrupt their life cycle, these natural insecticidal substances have received renewed public, private, and regulatory interest for reducing the use of synthetically designed pesticides. This trend is most interesting since these natural chemicals or complexes of chemicals, when isolated and reproduced synthetically, tend to be also naturally biodegradable and thus extremely unlikely to cause hardship and damage to the nontarget populations, to the environment, or to the crops being protected. This observation rests on the obvious fact that the environments where particular sets of plants have evolved to their present condition of self-protection have been naturally recycling for aeons and yet these natural environments are quite lush, friendly, and habitable. And they surely do not resemble chemical waste dumps or polluted waterways.

Selecting Chemicals for Toxicologic Characterization

The prodigious numbers and types of chemicals available for study and the length of time and cost of conducting adequate and acceptable carcinogenesis experiments in rodents (13) makes more difficult and more important the effort to nominate, select, and characterize toxicologically those chemicals that will allow us to make the best decisions to better protect public health. Chemicals selected for carcinogenesis evaluation are chosen primarily on the predilection that they might reasonably be carcinogenic. Some of the evidence may come from positive mutagenicity or, in particular, clastogenicity effects and structural features of probable carcinogenicity. Of equal importance to the selection process are the patterns and evidence of substantial human exposure, the production levels and distribution modes, and at times the interests and needs expressed by governmental research and regulatory agencies.

Chemicals are also selected through international cooperative efforts to share not only the costs of conducting experiments but to agree on a common core protocol to better allow for universal comparability (14). However, the protocol could be supplemented or enhanced with designs that would permit scientific answers or leads to other basic or applied research interests. Such designs might easily include the following:

1. 'Stop-start' exposure studies to evaluate progression/regression of induced lesions and tumors;
2. Timed evaluations of cellular proliferation/cellular turnover rates in extra groups of animals to identify the impact of this biologic response or lack thereof on the multistep carcinogenic process. The measurement of DNA damage and repair as well as adduct formation;
3. Intermittent exposure patterns to mimic certain industrial or home exposures (methylene chloride exposure for three days each week as might be typical for hobbyists);
4. Pharmacokinetics (or metabolism) studies to investigate the effects of long-term exposures and ageing on this parameter;

5. Sex- and species-specific experiments on a particular chemical, with each active researcher conducting one part of the multipart study, particularly for later comparisons.

The National Institute of Environmental Health Sciences (NIEHS) contingent to this cooperative effort nominated ten chemicals in rank order of importance for long-term studies: ethylene, ethyl benzene, methanol, carbon disulfide, 2,4-dichlorophenoxyacetic acid (2,4-D), n-hexane, aspartame, phenol, dimethylformamide, and 1-naphthylamine. On a more global basis the International Agency for Research on Cancer (IARC) periodically holds ad hoc workgroups to nominate chemicals for carcinogenesis evaluation, as well as to select chemicals as subjects for future IARC monographs.

Laboratory Resources

Worldwide capacity for conducting these essential evaluations is limited. An international survey reports (15) that 88 organizations located in 20 countries can undertake long-term chemical carcinogenesis studies on about 1000–1200 chemicals. Thus, approximately 300 additional chemical studies can be started each year, since one study takes five or more years to design, conduct, gather and tabulate data (histopathology for instance takes one pathologist roughly 1000–1500 hours—six to nine months—to complete a study involving two species, both sexes, three exposure groups and controls each comprised of 50 animals), evaluate and interpret the results, and prepare the report. Total capacity would presumably increase if more money were available to do these experiments, although money alone will not automatically increase the number of chemicals studied. Other factors are involved: the limited supply of knowledgeable staff and adequate facilities; reluctance or inability to conduct these semi-routine studies; lack of competence or qualifications among those individuals or laboratories actively seeking to do these time-consuming, meticulous, and laborious investigations. Thus the search continues for shorter-term, less complicated, more easily performed, and cheaper in vitro or in vivo assays (and perhaps eventually computer-simulated model experiments). However, most scientists in these fields of research believe that the day for replacing, or even reducing, current efforts or for duplicating the value of long-term in vivo carcinogenesis studies will be far into the future.

Currently 10 of the 20 countries are responsible for more than 95% of these studies, and four countries conduct 87% of those: Federal Republic of Germany [195], France [14], Hungary [27], India [18], Italy [11], Japan [174], The Netherlands [37], Norway [11], United Kingdom [67], and United States [609]. In the US for example, the NTP is responsible for about 50% of the 609 chemicals being evaluated, and in Japan the National Institute of Hygienic

Sciences is conducting nearly 31% of those 174 being examined. Estimates on how many chemicals have been studied for potential carcinogenicity generally do not consider whether the study was adequate or not, and frequently the tabulations include questionable assays such as cell tranformation and chromosomal aberration systems. The IARC Monographs contain their carcinogenicity evaluations on around 800–900 chemicals, groups of chemicals, mixtures, and occupations (16).

EXPERIMENTAL PROTOCOL and DESIGN

> One must have mammalian studies and those of the rat, the ape, and the human are three of the most important (17).

Carcinogenesis studies are formulated to include pertinent chemical-specific biological properties into the study design (18). These toxicology studies are typically carried out using both sexes of two species of rodents divided randomly into sets of 50–60 animals per exposure and control groups; three exposure concentrations are graduated down from a top level, a level of exposure carefully and judiciously chosen to show some minimal yet obvious chemically associated toxicity of a degree that should not unduly compromise the normal well-being or growth and survival of the animals. The species most often used by the NTP are the inbred Fischer 344 rat and the hybrid B6C3F1 (C3H × C57B16) mouse. Duration of exposure is generally two years (or about two thirds of the life span of these rodent species). Single, intermittent, or varied exposures may be used to mimic specific occupational situations. Exposures can be started prior to conception, during gestation and lactation, and for specified times thereafter.

Some investigators or organizations use protocols that expose animals for 18 months (generally short-lived species and strains), whereas others extend to 30 or 36 months, or to the full life span (usually until a preselected small percentage of animals remain, e.g. 10% or 20%). Animals are assessed for visible lesions during necropsy and a prescribed list of tissues and organs are taken for microscopic evaluation. These diagnoses are substantiated and subject to peer review. The data are tabulated and statistical comparisons are made that adjust for possible differences in survival between groups. The collated findings are evaluated, interpreted, and presented in public meetings to a nongovernmental peer review panel of experts in chemical carcinogenesis.

Special studies added to the core or typical study design might include: metabolism, pharmacokinetics, reproductive- and immunotoxicology, stop exposure groups, cell proliferation, interim evaluations (at 12, 15, 18, or 21 months) to observe toxic and early preneoplastic or neoplastic lesions, oncogene and suppressor gene expressions, DNA damage and repair and adduct formation, and other paradigms as identified. Such supplemental studies help to better define the long-neglected influences of age and stress on the routine processes of metabolism, cell proliferation or cell turnover, tumor progression or tumor regression, and to provide greater insight into potential mechanisms or understanding of the various species, strain, sex, and organ differences in response to chemical carcinogens.

Until recently the impact of various single or multiple parameters on the long-term process of chemical carcinogenesis received scant attention. Much needs to be learned about these essential but unknown factors. Although the findings from these experiments are indeed important and valuable for identifying potential long-term hazards to humans, possible effects in humans could be predicted more reliably, and better animal models selected if, for example, an extensive catalogue existed on normal and exogenously induced cell turnover and cell proliferation patterns as well as pharmacokinetic data to better allow evaluation of scientifically based hypotheses.

Each study on a particular chemical usually involves four individual, separate yet concurrent, experiments: male rats, female rats, male mice, female mice. Ordinarily, groups of 50 animals (or more depending on supplemental studies) of each sex, species, and exposure level (including controls) are given the chemical by a single route of exposure: usually in the feed (in 60% of the 400 studies we evaluated), by oral intubation (27%), by inhalation (5%), by application to the skin (3%), by intraperitoneal injection (3%), or in drinking water (2%).

Routes of Exposure

The optimal route should most closely mimic the major human exposure route where possible. Microencapsulation, for example, is a newly developed technique for administering unstable or reactive agents in feed, mainly to replace gavage (19). Moreover, a policy about using single routes of exposure (or even single chemicals) should be broadened to more closely mimic the real human situation and also to discover the biologic effects of simultaneous mixed or multiple exposure routes in our experiments.

As an illustration, if humans are exposed to an agent by multiple and obvious routes, such as to trichloroethylene in water through bathing or showering or washing in general, through volatization from standing waters in closed spaces (toilets and cooking), or through drinking water, then animals should also be exposed by inhalation (to mimic vapor exposure), by dermal exposure (baths and showers), and orally (via drinking water or feed). Thus, the same animals would be exposed to a chemical or mixture of chemicals by at least three routes (in this example), in closer replication of the human exposure pattern. Although this procedure could prove complicated for determining the long-term total tolerated exposure, it would yield interesting results and should in fact become quite routine using the typical shorter-termed studies as multiple exposure route trials. Results could then be compared from single and multiple routes to administer the same chemical or mixture of chemicals. The physical and chemical properties of certain chemicals or classes of chemicals will preclude their use in these multiple-route experiments. This factor may strongly influence the selection of a route of exposure. Another novel approach would use a different chemical for each of several routes of exposure to ascertain not only synergism or antagonism, but also to learn the actual biological effects of real exposure patterns and situations for humans.

Exposure Level Selection Criteria

One of the most important and controversial issues surrounding long-term chemical carcinogenesis studies concerns not only the scientific criteria by which exposure concentrations are selected but the philosophy of toxicology as well. A prerequisite for toxicologic or other studies is that some chemically associated effects must be observed, otherwise the particular investigation will be considered wasteful of time and resources and of questionable value. A major challenge in designing long-term (i.e., two-years long) toxicologic experiments is to calibrate exposure levels to allow a reasonably normal laboratory life (health, appearance, body weight, etc) for the animals while guaranteeing obvious evidence of chronic toxicity over and above that typically seen in aged animals. This objective is exceedingly difficult to reach, especially since exposure concentrations are chosen from results obtained from relatively short exposure periods (ordinarily 13-week experiments) in robust animals. Rarely in practice is the optimal exposure level precisely identified; routinely this concentration remains unknown until the studies are completed. This retrospective issue becomes particularly important for hazard identification if in a positive carcinogenicity study the exposure levels are considered far in excess of those needed to maintain proper health and longevity of the animals and are clearly associated with the carcinogenic effects. Even under such circumstances, however, a positive response can never be discounted without further repeating the experiments or developing alternative explanations supported by strong scientific evidence; to do otherwise would not be in the best interests of public health. Of course, a 'no evidence' result in a scenario of low weight gain and poor survival or when the exposure levels had minimal or no effect would likely be considered an inadequate experiment that should be repeated.

A case in point is the evaluative history of 1,3-butadiene: the initial carcinogenicity studies showed an obvious carcinogenic response at exposure levels close to the Occupational Safety and Health Administration (OSHA) standard of 1000 ppm, yet both exposure groups exhibited such poor survival that the experiments had to be terminated unexpectedly at 60–61 weeks (20). If this had been a 'no evidence' study, the NTP would have had to repeat the studies at lower exposures. In fact, the results were overwhelmingly positive, and NIOSH called for a re-evaluation of the occupational standard. Additional studies of this high-volume industrial chemical were designed and conducted because the initial results were confounded at the exposures used (625 and 1250 ppm). Survival was reduced considerably due almost entirely to the carcinogenic effects, and neither was a dose response observed (both exposures caused similar effects) nor was it possible to differentiate among the many tumor types with respect to competing cancer risks. The latter unknown was eventually deciphered by an expanded experimental design that permitted improved survival at lower exposures (down to 6.25 ppm); when lower exposure levels obviated the competing risks of dying from hemangiosarcoma of the heart and lymphoma, dose-responses for different tumor types (e.g. lung) could be detected at exposures down to the lowest used (21), 1/100 of the exposure level shown to be carcinogenic in the earlier study. These experimental carcinogenicity data as well as results from occupational epidemiology studies (showing a positive association between exposure and hematopoietic and lymphatic cancer in humans (22, 23)) were used by OSHA to recommend that the permitted occupational exposure level be reduced from 1000 ppm to 2 ppm (a 500-fold reduction factor (24)). Here is another case whereby the exposures were lower than those measured in the occupational environment, and 1,3-butadiene must now be added to the list of human carcinogens that were first shown to cause cancers in laboratory animals.

Some of the criteria and information gained from the shorter-term studies currently used to select and establish appropriate and optimum long-term exposure concentrations are: (*a*) food and water consumption; (*b*) body weights; (*c*) clinical, pharmacologic, and toxicologic signs; (*d*) morbidity and mortality; (*e*) gross pathology; (*f*) organ weights; (*g*) histopathologic changes; (*h*) hematology, clinical chemistry, and biochemical parameters; (*i*) metabolism and pharmacokinetics; (*j*) and other endpoints, as appropriate (13, 18). Everyone who evaluates and interprets these studies should be familiar with the criteria and have experience in conducting these studies. Based mainly on the shorter-term experiments, in essence, exposure levels are chosen so as not to unduly compromise reasonable health and longevity, and to show some real toxicological indications of long-term chemical exposure.

LEVELS OF EVIDENCE OF CARCINOGENICITY

Carcinogens are those substances which produce a significant increase in tumor incidence when administered at any dosage level by any route of administration in any species of animal as compared to [concurrent] controls (25).

The five levels of evidence of carcinogenicity created and established for the National Toxicology Program were formulated because, in our view, simply labeling a chemical as either *carcinogenic* or *not carcinogenic* was not in the best interest of science or public health and was potentially confusing to administrators and the regulatory agencies in the United States and in other countries. Such a black and white appraisal system is problematical in that it does not allow for any scientific nuances that would permit a more realistic awareness of what the data and results actually mean. Although changing to a system of five categories may not appear to be substantive, we believed then and continue to believe that our five levels of evidence do permit flexibility and better differentiation. Despite initial objections that the proposal was too far-reaching and potentially confusing, most counselors and colleagues endorsed our efforts, and time has shown considerable benefits over the older system.

In the late 1970s, the International Agency for Research on Cancer (IARC) created a system of categories of evidence of carcinogenicity to address the need for more scientifically based definitions as well more clarity and ease of recognition. The IARC system of a balanced and objective evaluation of experimental and human data is without peer, but any categorical system has its frailties. Therefore, ad hoc Scientific Workgroups have been periodically convened to scrutinize, re-evaluate, and revise and rewrite their guidelines to keep pace with scientific development; the latest gathering was held in 1987 (26).

The most significant differences between the IARC and the NTP carcinogenicity programs and the major attribute of the IARC system is that they give single and comprehensive evaluations to all relevant and available experimental data, to all published epidemiological information, and to all experimental and epidemiological data combined. This innovation is a major contribution to public health.

NTP Levels of Evidence of Carcinogenicity

The immediate definitional distinction for NTP is that the data we evaluate come from our own experiments; that is, we design the experiments, oversee the conduct of these studies, validate the results, evaluate the data, present in public meetings the findings and conclusions before nongovernmental panels of experts, print the relevant experimental data and results in a series of technical reports, and publish the essence in scientific journals (27). Another unique feature of the Program stems from our decision not to make overall or combinational evaluations for a chemical; that is, we evaluate and separately report levels of evidence for each experimental grouping or cohort: male rats, female rats, male mice, female mice is the usual makeup of the experimental cells for each chemical study. As discussed later, however, comparing results across the four experimental groups may weaken or strengthen a particular level of evidence.

The initial levels of evidence were first described and used at a public peer review meeting held at the National Institute of Environmental Health Sciences in June 1983. Five categories of interpretative conclusions were adopted for the Program's technical reports series to better maintain consistency with time and to better portray the actual 'strength' of the experimental evidence (and not to define potency or mechanism); Clear Evidence of Carcinogenicity; Some Evidence of Carcinogenicity, Equivocal Evidence of Carcinogenicity, No Evidence of Carcinogenicity, and Inadequate Study of Carcinogenicity. Thus, there are two categories for positive results ('Clear Evidence' and 'Some Evidence'), one category for uncertain findings ('Equivocal Evidence'), one category of no observable effects ('No Evidence') [note: neither IARC nor NTP use the misleading term 'negative'], and one category for experiments that because of major flaws cannot be evaluated ('Inadequate Study'). These categories refer to the strength of the experimental evidence and not to either potency or mechanism. Moreover, from our perspective and for our public health purposes, these five levels seemed to fit all (or at least almost all) of the situations we have encountered or would encounter as we evaluate the findings and results from our experiments. Lastly, we concluded that to fine-tune the categories of 'positive' experiments into more than two would be more analytical or quantitative then the experiments or data would permit, and further might make later independent evaluations by others more difficult.

In March 1986, the levels of evidence were re-examined by another ad hoc group of experts, again in public session, and select changes were proposed and made to the definitions. The Program staff also agreed to use these newly defined and newly named levels of evidence for a period of time, and then to re-evaluate them at some other future public meeting. As currently used the levels and definitions are:

Clear evidence of carcinogenic activity describes studies that are interpreted as showing a dose-related (*a*) increase of malignant neoplasms, (*b*) increase of a combination of malignant and benign neoplasms [in the same organ], or (*c*) marked increase of benign neoplasms if there is an indication from this or other studies of the ability of such tumors to progress to malignancy.
Some evidence of carcinogenic activity describes studies that are interpreted as showing a chemically related increased incidence of neoplasms (malignant, benign, or combined) in which the strength of the response is less than that required for clear evidence.
Equivocal evidence of carcinogenic activity describes studies that are interpreted as showing a marginal increase of neoplasms that may be chemically related.
No evidence of carcinogenic activity describes studies that are interpreted as showing no chemically related increases in malignant or benign neoplasms [or a combination of malignant and benign neoplasms].

Inadequate study of carcinogenic activity describes studies that because of major qualitative or quantitative limitations cannot be interpreted as valid for showing either the presence or absence of carcinogenic activity.

OTHER CONSIDERATIONS The most significant benefit from the March 1986 meeting was a listing of those items that ordinarily would impact on the conclusion about a particular experiment or data set. Thus the Program staff subsequently compiled a collection of key factors that should be (and were) considered that would extend the actual boundary of an individual category of evidence. This compilation allows for incorporation of scientific experience and current understanding of long-term carcinogenesis studies in laboratory animals, especially for borderline evaluations between two adjacent levels of evidence. These considerations should include: (*a*) the adequacy of the experimental design and conduct; (*b*) occurrence of common versus uncommon neoplasia; (*c*) progression (or lack thereof) from benign to malignant neoplasia as well as from preneoplastic to neoplastic lesions; (*d*) the capacity of some benign neoplasms to regress whereas others (of the same morphologic type) progress. At present it is impossible to identify the difference. Therefore, where progression is known to be a possibility, the most prudent course is to assume that benign neoplasms of those types have the potential to become malignant; (*e*) combinations of benign and malignant tumor incidences known or thought to represent stages of progression in the same organ or tissue; (*f*) latency in tumor induction; (*g*) multiplicity in site-specific neoplasia; (*h*) metastases; (*i*) supporting information from proliferative lesions (hyperplasia) in the same site of neoplasia or in other experiments (same lesion in another sex or species); (*j*) the presence or absence of dose relationships; (*k*) the statistical significance of the observed tumor increase; (*l*) the concurrent control tumor incidence as well as the historical control rate and variability for a specific neoplasm: (*m*) survival-adjusted analyses and false positive or false negative concerns; (*n*) structure-activity correlations; and (*o*) in some cases, genetic toxicity. Of course more key factors could be added to this list, but these main factors should always be considered while evaluating the long-term chemical carcinogenesis data and results.

Mechanisms of action are not yet sufficiently understood to become significant factors in the evaluations on a global or generic basis, but certainly all relevant biologic information (oncogene activation and tumor suppression, pharmacology and pharmacokinetics, DNA damage and repair, among others) should be considered when deciding a single level of evidence of carcinogenicity. At present we are unable to predict the eventual impact of some of these pieces of information on the relative strength or weakness of a particular level of evidence. A most important criterion is knowledgeable and objective scientific staff. Individuals should have direct experience in the design, conduct, data collection, results interpretation, and evaluation of experiments.

Since their adoption by the National Toxicology Program in June 1983, the levels of evidence of carcinogenicity have been used to evaluate 120 long-term chemical carcinogenesis studies comprised of 450 individual experiments. The independent peer review panel concurred on 427 of these 450—for a 95% concurrence with the proposed evaluations made by Program staff. Roughly one half of these proposed changes were for decreases in the levels of evidence and one half were for increases in the levels of evidence. In no instances did the panel recommend a change of more than one level of evidence.

If the experiments were designed and conducted adequately and showed 'no evidence of carcinogenicity', in the judgment of the Program staff and subsequent ad hoc peer review panel, then these evaluations, in which the laboratory animals do not have a incidence of site-specific neoplasia greater than concurrent control animals, do not mean necessarily or absolutely that a chemical is not a carcinogen inasmuch as the experiments are conducted under a limited set of conditions. Nonetheless, the NIEHS/NTP endorses the concept that under the conditions of exposing both sexes of two rodent species to relatively high exposure levels for extended periods of times (at least two years), it is scientifically reasonable, for practical purposes, to regard those chemicals that did not induce neoplasia as presenting little potential carcinogenic hazards to humans. However, new data or scientific advances could later influence this interpretation.

Conversely, positive carcinogenicity results in these experiments demonstrate that a chemical is carcinogenic for laboratory animals under the conditions of the study and indicate that exposure to the chemical should be regarded for prudent public health and scientific purposes as being a likely carcinogenic hazard to humans. Of course, all the available and relevant scientific information on a particular chemical should be evaluated before changes in occupational or public health policies are proposed or initiated. Nonetheless, these findings would lead most reasonable people to conclude that exposures to such chemicals should be reduced, minimized, or eliminated where appropriate.

DATA FOR EVALUATION AND POLICY SETTING

> I have no data yet. It is a capital mistake to theorize before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts.
>
> The Adventures of Sherlock Holmes, A. Conan Doyle

Table 2 Summary results from 379 rodent carcinogenicity studies

Study result	By experiment: Male rats	Female rats	Male mice	Female mice	Total	By chemical: Rats	Mice	Overall
Positive	123	104	102	122	451	142	135	193
Equivocal	42	32	35	21	130	46	36	51
No evidence	181	213	201	202	797	163	176	130
Inadequate	15	13	15	8	51	11	6	5
Totals	361	362	353	353	1429	362	353	379

The National Cancer Institute (NCI) and the National Toxicology Program (NTP) have collectively carried out and evaluated 379 long-term carcinogenicity studies in rats and mice. These consisted of 1429 separate sex- and species-specific experiments, and the results are summarized in Table 2. Approximately one half—193/379, 51%—of the studies showed a carcinogenic effect in at least one organ in one of the four sex-species groups. In addition, 13% (51/379) produced equivocal findings, 34% (130/379) of the chemicals gave no evidence of a carcinogenic effect, and 1% (5/379) of the studies were considered inadequate for evaluation. All four sex-species groups appeared to have about equal overall sensitivity for detecting the various carcinogenic effects produced by the 379 chemicals evaluated. The proportion of positive experiments was 34% for male rats, 29% for female rats and for male mice, and 35% for female mice.

Table 3 summarizes the distribution of individual outcomes in the 311 chemical studies with adequate experiments in all four sex-species groups. (In this tabulation, equivocal results were placed into the no evidence category. The important message about equivocal findings is that they do show marginal or uncertain neoplastic findings that cannot be simply discounted, yet they do not fit easily into either the positive or no evidence categories). The most frequent outcomes were for chemicals considered to give no evidence of carcinogenicity in all four experimental groups (150/311 or 48%). From the public health point of view, perhaps the forty-two chemicals (actually forty, since ethylene dibromide (EDB) and dibromochloropropane (DBCP) were positive in all four groups in two separate sets of experiments using different routes of exposure), causing cancer in all four groups (42/311 or 14%). should be given the most immediate attention (Table 4). The overwhelming or unique carcinogenic responses for certain of these chemicals—tetranitromethane (which caused concentration-related lung neoplasms in each of the four experimental groups with no associated toxicity, and caused considerable nasal toxicity without neoplasia; this finding suggests

Table 3 Carcinogenicity results for 311 chemical studies in rodents

Proportion of positive studies	Male rats	Female rats	Male mice	Female mice	Number of studies with these results		%
4/4	+	+	+	+	42		
					Subtotals	42	13.5
3/4	+	+	+	−	1		
	+	+	−	+	11		
	+	−	+	+	7		
	−	+	+	+	6		
					Subtotals	25	8.0
2/4	+	+	−	−	19		
	+	−	+	−	2		
	+	−	−	+	7		
	−	+	+		2		
		+	−	+	3		
	−	−	+	+	25		
					Subtotals	58	18.7
1/4	+	−		−	17		
	−	+		−	4		
	−	−	+	−	8		
	−	−		+	7		
					Subtotals	36	11.6
0/4*	−	−	−	−	150		
					Subtotals	150	48.2
Totals						311	100

*Equivocal evidence (or marginal) results are considered to be between a positive response and no evidence of a response. For these data equivocal results were placed into the no evidence category.

that human cohorts in the munitions industry should be identified and investigated for lung cancer) and glycidol (which caused 7 to 10 chemically associated tumor increases in rats and mice)—might be used to stimulate industrial or governmental interest in defining occupational or environmental cohorts that could be investigated epidemiologically.

Results Correspondence

Correlations in carcinogenic response among the various sex-species groups confirm those reported previously (28). The concordance is particularly high for males and females within a species (Table 5): in 88% of the chemical studies in mice and in 85% of the experiments in rats the responses were either both positive for chemically induced cancers or both showed no evidence of

Table 4 Chemicals showing evidence of carcinogenicity in each of the four sex-species experimental groups

Chemical	Technical report no.	Routes of exposure
3-Amino-9-ethylcarbazole	93	Feed
o-Anisidine hydrochloride	89	Feed
Benzene	289	Gavage
Bromodichloromethane	321	Gavage
Chlorinated paraffins: C12, 58% chlorine	308	Gavage
3-Chloro-2-methylpropene	300	Gavage
4-Chloro-o-phenylenediamine	63	Feed
C.I. basic red 9 monohydrochloride	285	Feed
p-Cresidine	142	Feed
Cupferron	100	Feed
2,4-Diaminoanisole sulfate	84	Feed
1,2-Dibromo-3-chloropropane (DBCP)	28	Gavage
1,2-Dibromo-3-chloropropane (DBCP)	206	I.P. Injection
1,2-Dibromoethane (ethylene dibromide)	210	Inhalation
1,2-Dibromoethane (ethylene dibromide)	86	Gavage
1,2-Dichloroethane	55	Gavage
Dichloromethane (methylene chloride)	306	Inhalation
Di(2-ethylhexyl) phthalate	217	Feed
Diglycidyl resorcinol ether (DGRE)	257	Gavage
Dimethylvinyl chloride (DMVC)	316	Gavage
1,4-Dioxanc	80	Water
Ethyl acrylate	259	Gavage
Ethylene thiourea (ETU)	388	Feed
Glycidol	374	Gavage
4,4'-Methylenedianiline dihydrochloride	248	Water
2-Methyl-1-nitroanthraquinone	29	Feed
Michler's ketone	181	Feed
Nitrilotriacetic acid (NTA)	6	Feed
4,4'-Oxydianiline	205	Feed
Phenoxybenzamine hydrochloride	72	I.P. Injection
Polybrominated biphenyls (FF-2)	244	Gavage
Procarbazine hydrochloride	19	I.P. Injection
Propylene oxide	267	Inhalation
Sulfallate	115	Feed
2,3,7,8-Tetrachlorodibenzo-p-dioxin	209	Gavage
Tetrachloroethylene	311	Inhalation
Tetranitromethane	386	Inhalation
4,4'-Thiodianiline	47	Feed
o-Toluidine hydrochloride	153	Feed
Tris(1-aziridinyl)phosphine sulfide (THIO-TEPA)	58	I.P. Injection
Tris(2,3-dibromopropyl) phosphate	76	Feed
4-Vinyl-1-cyclohexene diepoxide	362	Skin paint

[Note: For several other chemicals evaluated by us in fewer than four experiments (e.g., one experiment may have been considered inadequate, or others reported to be adequate positive studies in another species, or the data were replicated in the same species/strain, and for completeness these are given here: 1,3-butadiene (2 of 2 in mice and replicated—TR No. 288, positive in rats reported); 1,3-dichloropropene (Telone II, 3 of 3 and the other experiment showed similar target sites—TR No. 269); ethylene oxide (2 of 2 in mice—TR No. 326, positive in rats reported); ochratoxin A (2 of 2 in rats—TR No 358, positive in mice reported); pentachlorophenol (2 of 2 in mice and replicated—TR No. 349)].

Table 5 Intra- and inter-species concordance in carcinogenic responses in 311 chemical carcinogenicity studies in rodents

Comparison	Observed response ++	+−	−+	−−	% Concordant (++ or −−) responses
Male rats vs female rats	85	37	16	206	84.6 (291/344)
Male rats vs male mice	53	54	41	167	69.8 (220/315)
Male rats vs female mice	69	41	42	167	74.0 (236/319)
Female rats vs male mice	52	38	43	183	74.4 (235/316)
Female rats vs female mice	64	28	48	181	76.3 (245/321)
Male mice vs female mice	89	13	29	205	87.5 (294/336)
Rats vs mice	81	40	40	150	74.3 (231/311)

carcinogenicity. For rats and mice overall, the corresponding concordance was 74%, indicating a respectable yet not perfect qualitative extrapolation index. Nonetheless, for this reason long-term carcinogenesis studies should continue to use at least two species. Yet because mimicry in tumor response among sexes within a species is 85% or better, then opposite sexes from each species could be used to allow for more studies on more chemicals at relatively lower costs. Using a male and a female still would allow for an evaluation of the influence of hormonal differences.

For the question whether results from animal experiments will be considered and accepted as valid and relevant to the human situation, a plethora of papers, articles, books, and symposium proceedings has addressed the issues of extrapolation—from individual to individual, from sex to sex, from strain to strain, from species to species, from race to race, from geographic location to geographic location—and no clear consensus of interpretation or harmonization of thought has been reached. Nevertheless, there is strong belief in the scientific community that experimental data and interpretative information obtained from whole animals are indeed relevant and applicable to the expected or, in several cases, observed responses in humans; this is especially true for chemically associated cancers in humans and in rodents (9, 29). Using our collective experience, we believe experiments using animal model surrogates for humans are the most reliable means for determining potential carcinogenic hazards to people exposed to chemical carcinogens. Good epidemiological studies are certainly needed to give us the best answers; yet the sensitivity of human studies is generally poor and most important studies have confounding flaws (lack of exposure data, small cohort populations, cumbersome follow-up, recall bias, and other retrospective difficulties) that make them less reliable than the long-term experiments in laboratory animals. Likewise, short-term in vitro or in vivo assays are not adequate substitutes for the long-term studies for many reasons: the full metabolic capability is not present; most chemicals need metabolic activation to an ultimate carcinogen; chronic exposures are usually not possible; the systems are somewhat 'artificial'; initiation is often accomplished with a potent and organ-specific carcinogen; long-term in vivo mammalian studies mimic closest overall human biology; and these assays have not been shown to be as reliable as results from the two-year studies for predicting effects in humans.

Protocol Modifications

As we continue to introduce appropriate and reliable refinements into these carcinogenesis assays (30), cumulative retrospective analyses clearly indicate that the use of only male rats and female mice (to cover the influence of gender-hormonal effects) would have allowed us to identify correctly and qualitatively 297 of the 311 chemicals (or 96%) evaluated for long-term carcinogenesis effects. Using the 161 chemicals showing evidence of cancer in at least one of the four experiments conducted on each chemical shows that for 14 chemicals (or 9%) the positive cell was in either the female rat, the male mouse, or both of these groups; and these chemicals would not have been identified as associated with the induction of any cancer if the 'reduced protocol' had been used instead of the more typical protocol of all four sex-species groups. As Table 3 shows, four chemicals showed carcinogenic effects only in female rats, eight showed effects only in male mice, and two chemicals showed carcinogenic effects in both female rats and in male mice. These chemicals with their chemically associated lesions are given in Table 6.

Additionally, the evidence of carcinogenicity for twelve other chemicals would have been reduced from two to one species (Table 7), which certainly influences whether or not the IARC or the NTP would list these chemicals as likely human carcinogens. Thus, a decision is needed on whether society would be well served by the use of a reduced protocol to evaluate close to twice as many chemicals as are now being studied. This decision raises serious policy questions for the regulatory agencies about the impact of such a reduced protocol on their individual and collective regulatory processes. On the other side, many investigators frequently use only one species, and others will use only one sex of one species; these experimental procedures appear to be quite acceptable if the chemical shows a carcinogenic effect or is being used for research purposes, but are quite unacceptable for showing no evidence of carcinogenicity and, hence, implied 'safety'.

Using our retrospective analyses (Table 8) as a prospective indicator of anticipated results from the use of a reduced protocol on the next 400 or so chemicals shows that for the combination of male rats and female mice we would be absolutely correct (+,+ and −,−) in identifying a two species carcinogen or no evidence chemical 62% (192/311) of the time. For a qualitative yes or no this combination would be 'correct' 86% of the time. For

Table 6 Using male rats and female mice would have "missed" the carcinogenic effects of these fourteen chemicals shown to be positive in 311 studies

Sex-species	Chemical	Salmonella	Site of Neoplasia
Female Rats	Daminozide	(−)	Uterus [mm: liver]
	p-Quinone dioxime	(+)	Urinary bladder
	Trimethylthiourea	(−)	Thyroid gland
	C.I. Acid orange 3	(+)	Kidney
Male Mice	Aldrin	(NT)	Liver [mr: thyroid gland] [fr: thyroid gland]
	Allyl glycidyl ether	(+)	Nasal passage [mr: nasal passage] [fm: nasal passage]
	3-Amino-4-ethoxyacetanilide	(+)	Thyroid gland
	Chlorinated paraffins (C23, 43% Cl)	(−)	Lymphoma [fr: pheochromocytomas] [fm: liver]
	C.I. Vat yellow 4	(−)	Lymphoma
	Dicofol	(−)	Liver
	3-Nitro-p-acetophenetide	(+)	Liver
	Piperonyl sulfoxide	(−)	Liver
Female Rats and Male Mice	Nithiazide	(+)	fr: Mammary gland mm: liver [fm: liver]
	Phenylbutazone	(−)	fr: kidney mm: liver [mr: kidney]

14/311 = 5%; [ ] = equivocal response

'positive' studies alone the predictive value would be 72% and for 'no evidence' chemicals the forecast would be 91%. Thus, 14 (or 9%) of the 'positives' would have been 'missed' using this combination: 2 female rat/male mouse positives, 4 female rat positives, and 8 male mouse positives. Seven of the two-species positives would have been reduced to a single-species positive. These retrospective facts are pertinent to a decision on the advisability of using (selectively) the modified protocol.

Perhaps a reduced protocol could be used only for 'structural certainties': benzidine dyes and derivatives, anthraquinones, nitrosamines, phenylenediamines, anilines. Or should we assume with some degree of scientific confidence that chemicals in these structural classes would be carcinogenic, and not even bother with a confirmatory long-term study?

Another refinement that would substantially enhance our ability to associ-

Table 7 Using male rats and female mice would have "reduced" these 12 chemicals to a single species positive

Chemical	Salmonella	Male rats	Female rats	Male mice	Female mice
Benzofuran	(−)	NE	SE	CE	CE
Chlorendic acid	(−)	CE	CE	CE	NE
p-Chloroaniline HCl	(+)	CE	EE	SE	NE
HC Blue 1	(+)	EE	SE	CE	CE
Hexachlorodibenzodioxin	(−)	E	P	P	P
ICRF-159	(NT)	N	P	N	P
Isophosphamide	(NT)	N	P	N	P
1,5-Naphthalenediamine	(+)	N	P	P	P
Nitrofurazone	(+)	EE	CE	NE	CE
p-Nitrosodiphenylamine	(−)	P	N	P	N
Phenesterin	(−)	N	P	P	P
Tetrachlorovinphos	(−)	N	P	P	P

P = Positive; E = Equivocal; N = Negative; CE = Clear evidence; SE = Some evidence; EE = Equivocal evidence; NE = No evidence

Table 8 Retrospective analyses as a prospective indication of anticipated results

Carcinogenic response If "True"	Then "Predict" FR	MM	No. of chemicals	%	Agree Qualitative Yes	or No
For 67 chemicals					60	7
if	+	+	42	63		
MR + &	+	−	11	16		
FM +		+	7	10		
then			7	10		
For 39 chemicals					22	17
if	+	+	1	3		
MR + &	+	−	19	49		
FM −	−	+	2	5		
then			17	44		
For 41 chemicals					34	7
if	+	+	6	15		
MR − &	+		3	7		
FM +	−	+	25	61		
then	−	−	7	17		
For 164 chemicals					150	14
if MR −	+	+	2	1		
& FM −	+	−	4	2		
then	−	+	8	5		
	−	−	150	91		
		Total	311		266 (86%)	45 (14%)

ate or to discount single-site or single-organ carcinogenic effects would be to do step-sections or serial-sections on the particular organ showing evidence of carcinogenicity. This could be made routine to the protocol, or as a supplement in those cases where the increases are marginal and one wants to establish the actual incidence as closely as possible. Using the kidney as an example, most investigators performing long-term studies routinely take single (or sometimes double) sections of each kidney, and sections where gross lesions are visible. In our experience, one cannot predict whether the evidence will be enhanced or decreased by evaluating more sections of kidney. A statistical or logical hypothesis might predict that more lesions would be found in direct proportion to the number of microscopic sections evaluated; however this was not the case in the limited number of instances we investigated. This finding should not be surprising because "certainty is unattainable in science" (31), and constraint is appropriate when formulating biological hypotheses since biology (especially in the absence of data) is far from being as congenial and as predictive as we would like to believe.

For several experiments using different chemicals there were increases—some marginal—in tumors of the kidney. To further verify these carcinogenic responses, as many additional histological sections of kidney as was practicable were made from the remaining kidney tissue: unfortunately, only six to ten sections were possible. This procedure was followed for five chemicals that showed evidence of effects in the kidney; the additional sectioning gave some interesting but not unexpected results (i.e., no consistency): for two of the chemicals evidence for an association between chemical exposure and kidney tumors increased significantly, for two others an association was confirmed, and for the remaining chemical evidence of an association decreased. The impact of multisections on the incidence of tubular cell tumors in the five concurrent control groups of male rats showed an overall increase from an historical background rate of 0.5% to around 5%, or a tenfold increase; this must be tempered somewhat because two of the control groups showed 8 of the 10 additional tumors among the 210 control animals. For the few additional studies on control female rats the consistent lack of background tumors of the kidney was confirmed.

As an example of a different target site, a chemical that showed a possible association with the occurrence of brain tumors was sustained after evaluating several additional sections. Thus, for single-site carcinogens there seems to be ample evidence for additional histopathology to confirm or to contradict a possible association between chemical exposure and cancer.

Importantly, caution is needed when comparing contemporary cancer incidences in exposed groups with those seen in historical groups since trends change over time and the causes of 'spontaneous' tumors are unknown. As stated by Boyland in 1969 (32), "Cancer, like other natural phenomena, has causes. When a tumor is described as being spontaneous it means that the causes are unknown, like those of most events which occur in living things."

Target Organs

For further species correlations and extrapolation paradigms, the most frequent and common tissue or organ sites of chemically induced carcinogenicity and of toxicity offer other useful indicators of comparison. Table 9 lists the ten most frequent sites of carcinogenicity for both rats and mice. The liver is the most frequent site of carcinogenicity in both species; liver tumors appear to be relatively more common in mice than in rats. The historical control incidence for liver tumors in NTP studies is 3% for male F344 rats, 1% for female F344 rats, 35% for male B6C3F1 mice, and 13% for female B6C3F1 mice. Thus, while liver tumors are twice as frequently induced in mice than in rats, the liver appears to some extent to be responsive to carcinogenic effects independently of the background rate or species. For those chemicals causing liver tumors in these studies, many induced these neoplasms in both species or in both male and female mice. Also, most increased responses were seen at both exposure levels and liver carcinomas predominated (33). Furthermore, the majority of the chemicals studied are not associated only with the induction of mouse liver tumors, and of the chemicals that do cause neoplasia in rodents, most that give positive responses cause neoplasms in other organs as well. Pharmacologically, it is not surprising that the liver should be involved in these long-term studies simply because it is the primary organ of metabolic activation and degradation of chemicals. We agree with the concept that "when used conservatively and responsibly, lesions of the liver are our single most valuable tool in evaluating the biologic effects of chemicals" (34).

The relevance of liver lesions in mice has been considerably strengthened by data showing that oncogene patterns in chemically exposed animals are different from those from 'normal' liver tissues in exposed mice or from control mice. These proto-oncogenes are cellular genes that are expressed during normal growth and development processes, and can be activated to cancer-causing oncogenes by point mutations or by gross DNA rearrangement (chromosomal translocation or gene amplification) (35). These lesions are especially revealing for chemicals that are apparently nonmutagenic and yet cause point mutations in chemically (furan and furfural) exposed B6C3F1 mice (36). Distinct oncogene activation in spontaneous versus chemically induced (37, 38) and in benign versus malignant neoplasms (37, 39) greatly enhances the use of molecular events in the risk assessment process. Moreover, loss of specific regulatory functions (i.e., tumor suppressor genes) represents an important feature in neoplastic transformation (40–42). This finding is relevant because the H-*ras* gene has the same amino acid sequence in both humans and rodents (43).

Table 9 Most frequent sites of chemically induced carcinogenic effects in rats and mice

Rats		Mice		Rats and mice	
Sites	No. positive studies	Sites	No. positive studies	Sites	No. positive studies
Liver	41	Liver	84	Liver	100
Kidney	27	Lung	22	Hematopoietic system	28
Mammary gland	22	Forestomach	17	Lung	28
Thyroid gland	17	Hematopoietic system	15	Kidney	27
Zymbal gland	16	Circulatory system	11	Mammary gland	27
Hematopoietic system	16	Thyroid gland	9	Forestomach	23
Forestomach	15	Mammary gland	9	Thyroid gland	19
Urinary bladder	14	Ovary	8	Urinary bladder	16
Skin	13	Harderian gland	7	Zymbal gland	16
Clitoral/Preputial gland	9	Uterus	7	Uterus	15

In evaluating lung tumors from certain species, You et al (44) have identified what clearly appears to be the same activated K-*ras* oncogene in mouse and human lung tumors. Similarly, tumor suppressor genes have been identified across species (42,45). In basic cellular functions, *ras* genes are likely to play a fundamental role based on their high degree of conservation throughout eukaryotic evolution.

Other common sites of carcinogenicity in both species are forestomach, hematopoietic system, thyroid gland, and mammary gland. In contrast, other sites of carcinogenicity appear to be species- or sex-specific. For example, the kidney was the site of carcinogenicity in 24 male rat studies, whereas the corresponding numbers for female rats, male mice, and female mice were 9, 3, and 1, respectively (46). Interestingly, cancers of the kidney in humans are more frequently seen in males (11.1 per 100,000) than in females (5.2 per 100,000). Mesothelioma was increased in seven male rat studies, but in only one female rat study and in none of the studies in mice.

Some sites seemed to be affected more frequently in one species than the other. Examples include ovary (eight chemicals produced ovarian tumors in mice, while none did in rats), harderian gland (seven chemicals produced harderian gland tumors in mice, while none did in rats), zymbal gland (sixteen chemicals produced zymbal gland tumors in rats, while only two did in mice) and pancreas (eight chemicals produced pancreatic tumors in rats, while none did in mice). The sex- or species-related responses observed in these studies should not be thought of as flaws or confounding factors in the animal models, but are either simply associated with the different hormonal milieux between the sexes (apart from ovaries and uteri, we often observe lesions in comparable sex organs such as clitoral glands and preputial glands), or the mechanistic rationale for these differences has not yet been deciphered. Moreover, we witness similar sex differences in tumor patterns in the human population for both incidence rates and mortality rates (Table 10). There is remarkable correspondence for several of the top ten tumors sites seen most frequently in rodents (Table 9) and in humans (Table 10). These differences are being actively explored by various investigators in other disciplines as well as in experimental carcinogenesis, and perhaps some better explanations will be forthcoming. Until that time one should recognize that the same chemical may cause tumors at different sites and with different patterns between species and between sexes. This finding should not be taken to mean that the effects are not real or are not relevant to the process of extrapolation.

SUMMARY COMMENTS AND CONCLUSIONS

> The chemical environment of the present century whether in atmospheric pollution, in industrial operations, or in products sold for personal consumption, has greatly increased public health hazards (47).

Table 10 Age-adjusted human top-ten cancer rates (per 100,000)

Males and Females Site	Rate	Males Site	Rate	Females Site	Rate
		Cancer incidence rates			
Breast	56.9	Prostate gland	87.8	Breast	103.8
Lung and bronchus	55.9	Lung and bronchus	82.7	Colon/rectum	43.9
Colon/rectum	51.5	Colon/rectum	51.5	Lung and bronchus	35.7
Prostate gland	35.8	Urinary bladder	29.1	Cervic uteri Corp and Uter, NOS	30.2
Urinary bladder	16.7	Oral and pharynx	16.5	Ovary	13.4
Cervix uteri corp and uter, NOS	16.4	Non-Hodgkin's lymphoma	15.1	Non-Hodgkin's lymphoma	10.5
Non-Hodgkin's lymphoma	12.6	Leukemia	12.8	Melanoma of skin	8.7
Oral and pharynx	11.0	Stomach	11.9	Pancreas	8.2
Leukemia	9.7	Melanoma of skin	11.8	Urinary bladder	7.7
Pancreas	9.4	Kidney and renal pelvis	11.1	Leukemia	7.3
		Cancer mortality rate			
Lung and bronchus	46.7	Lung and bronchus	74.0	Breast	27.3
Colon/rectum	20.42	Colon/rectum	24.8	Lung and bronchus	26.7
Breast	15.3	Prostate gland	23.8	Colon/rectum	17.3
Prostate gland	9.1	Pancreas	10.1	Ovary	7.7
Pancreas	8.45	Leukemia	8.5	Pancreas	7.2
Leukemia	6.4	Stomach	7.4	Cervix uteri Corp and Uter, NOS	6.9
Non-Hodgkin's lymphoma	5.8	Non-Hodgkin's lymphoma	7.1	Leukemia	5.0
Stomach	5.0	Urinary bladder	5.9	Non-Hodgkin's lymphoma	4.7
Ovary	4.3	Esophagus	5.7	Stomach	3.3
Brain and nervous system	4.0	Oral and pharynx; Brain and nervous system	4.9	Brain and nervous system	3.3

Cancer Statistics Review 1973–1986, National Cancer Institute, May 1989

The fresh kills land fill on Staten Island discharges 4 million gallons of toxic leachate per day (48).

If cancer can be defined simply as a rebellion in an otherwise orderly cellular society (49), and if the multistage process of carcinogenesis is virtually consistent within the mammalian genre, then scientific logic would permit the obvious: chemically induced cancers would likewise be similar or extrapolative among the class mammalia. This is of course exactly what happens, based largely on our knowledge of the biology of tumor induction and progression (50, 51) and on the clear correspondence for those chemicals judged to be carcinogenic in humans that are also invariably carcinogenic in laboratory animals (8). Social and political debate begins when chemicals (especially those with economic importance) are shown unequivocally to induce cancer in laboratory animals and no relevant or reliable human data exist to confirm or to counter any association. In rare instances does acrimonious debate follow the demonstration in well-designed and properly conducted laboratory studies that a particular chemical is not carcinogenic. Differences in scientific opinion where they occur generally rest on interpretation of the data, on the system or model used to generate the data, on the person or organization conducting the investigations, and/or on the eventual impact these findings will have on our personal environments. As long as these findings are used to stimulate the regulatory process, political, social, and scientific debate will continue, almost regardless of the actual facts. In general, this dichotomy of thought usually benefits both sides of a particular issue, and often results in scientifically based compromise and a more centrist consensus.

Basic and applied toxicologic research results are equally important to confirm or substantiate equivocal epidemiological findings (or to further validate the current animal models as sentinel surrogates for humans). Conversely, chemicals considered to be potent or multiorgan carcinogens should serve as a strong signal for epidemiological studies. Further, obvious negative results from animal experiments can be used to discriminate among various chemicals or the typical human situation of multiple chemical exposures that are associated with known human toxicity or cancers. Consider an occupational setting where workers are exposed to these four high-volume solvents: benzene, methylene chloride, toluene, and xylenes. As we now know, the first is a potent animal carcinogen (and a known human carcinogen), the second is a multisite carcinogen studied at exposure concentrations not different than those monitored in furniture dipping and paint-stripping operations, while the latter two showed no evidence of chemically induced cancers in long-term, high exposure level studies. Thus, a prudent person would be advised to pick a solvent that did not show any evidence of carcinogenicity in experimental systems.

In our collective view, an overwhelming amount of scientific information better supports the value and use of animals as valid surrogates for humans than counters this evolutionary biologic attachment. If for the moment this could be taken to enjoin universal acceptance and agreement, enough unknowns or unproven hypotheses in chemical carcinogenesis would still keep us all happily busy for the rest of our lives in the quest for the real truth, or truths. The most prudent (and at times possibly least popular) public health course of action to take would be to reduce exposures to all manufactured chemicals to the absolute minimum that would still permit use of these agents that have become entrenched in our modern way of life. This may or may not be feasible, yet not to attempt to reduce the astronomical amounts of chemical releases into the environment would be in contravention of public health. Of course, one might concentrate first on those chemicals that are known to cause cancer in humans or in animals. As the IARC has recommended, "In the absence of adequate data on humans, it is biologically plausible and prudent to regard agents and mixtures [and exposure circumstances] for which there is sufficient evidence of carcinogenicity in experimental animals as if they presented a carcinogenic risk to humans" (16, 26).

We believe our scientific and public responsibility must continue to be directed toward identifying those chemicals, mixtures of chemicals, and exposure circumstances that present potentially the most predictable carcinogenic (and other toxicologic) hazards to humans. This we will continue to do using the best scientific advances available. The regulators and politicians, using the most knowledgeable science advisors (52 for instance), must balance the benefits and the known or predicted risks to human health before making difficult and public-oriented decisions. As Doll observed (53), "The final number of proven occupational carcinogens may, therefore, eventually be quite large."

Chemical carcinogenesis experiments are conducted primarily to identify those chemicals, mixtures of chemicals, or environmental and occupational exposures that may potentially induce cancer in humans. Besides human experience and epidemiological investigations, long-term studies in laboratory animals are the only validated and universally accepted means to determine carcinogenic hazards to the public health and to the environment. The important issue is not whether we are undeniably correct in extrapolating carcinogenic responses in laboratory animals up the evolutionary trail to humans, but rather to concentrate our vigilance to assure and improve the entrusted public health.

ACKNOWLEDGMENTS

We thank our colleagues Drs. Ron Melnick, Chris Portier, and John Bucher for their valuable scientific insights on the issues discussed in our paper, as well as for their friendly demeanor in meeting our requests for rapid reviews. We appreciate the help of Ms Donna Meyer in the production of this manuscript and the flexibility of Production Editor Ms Hilda Gould.

Literature Cited

1. Natl Res. Counc./Natl Acad. Sci. (NRC/NAS). 1977. *Drinking Water and Health,* p. 53. Washington, DC: Natl. Acad. Sci. 939 pp.
2. Browning, E. 1965. *Toxicity and Metabolism of Industrial Solvents.* Amsterdam: Elsevier
3. US Int. Trade Comm. (USITC). 1989. *Synthetic organic chemicals. United States production and sales, 1988.* USITC Publ. No. 2219
4. Natl. Res. Counc. 1984. *Toxicity Testing. Strategies to Determine Needs and Priorities.* Washington, DC: Natl Acad. Press. 382 pp.
5. Wanser, S. 1989. Pesticides and related products. See Ref. 3, Sect. 13:1–10
6. Thelin, A. 1990. Epilogue: Agricultural occupational and environmental health policy strategies for the future. *Am. J. Ind. Med.* 18:523–26
7. Huff, J. E., Haseman, J. K. 1991. Agrichemical pesticides and cancer. *Chem. Eng. News.* Jan. 7:33–36
8. Huff, J. E. 1991. Pesticides and human health. Identifying potential human health hazards from pesticides using carcinogenicity findings from experiments in laboratory animals. *Rev. Environ. Contam. Toxicol.* In press
9. Tomatis, L., Aitio, A., Wilbourn, J., Shuker, L. 1989. Human carcinogens so far identified. *Jpn. J. Cancer Res.* 80: 795–807
10. Elias, P. 1973. Editorial. *Toxicology* 1:1
11. Galloway, J. N., Thornton, J. D., Norton, S. A., Volchok, H. L., McClean, R. A. N. 1982. Trace metals in atmospheric deposition: A review and assessment. *Atmos. Environ.* 16:1677–1700
12. Arnason, J. T., Philogene, B. J. R., Morand, P., eds. 1989. *Pesticides of Plant Origin. ACS Symp. Ser.* 387. Washington, DC: Am. Chem. Soc. 213 pp.
13. Huff, J. E., McConnell, E. E., Haseman, J. K., Boorman, G. A., Eustis, S. L., et al. 1988. Carcinogenesis studies: results from 398 experiments on 104 chemicals from the U.S. National Toxicology Program. *Ann. NY Acad. Sci.* 534:1–30
14. Damstra, T., Kurokawa, Y. 1990. U.S.-Japan joint meeting on the toxicological characterization of environmental chemicals of mutual interest. *Environ. Health Perspect.* 87:301–7
15. IARC. 1988. *Information Bulletin on the Survey of Chemicals Being Tested for Carcinogenicity,* No. 13. Lyon, France: IARC. 391 pp.
16. IARC. 1972–1990. *IARC Monographs on the Evaluation of Carcinogenic Risk to Humans,* Vols. 1–49. Lyon, France: IARC
17. Koshland, D. E. 1988. Biological systems. Editorial. *Science* 240:1385
18. Chhabra, R. S., Huff, J. E., Schwetz, B. S., Selkirk, J. 1990. An overview of prechronic and chronic toxicity/carcinogenicity experimental study designs and criteria used by the National Toxicology Program. *Environ. Health Perspect.* 86:313–21
19. Melnick, R. L., Jameson, C. W., Goehl, T. J., Kuhn, G. O. 1987. Application of microencapsulation for toxicology studies. I. Principles and stabilization of trichloroethylene in gelatin-sorbitol microcapsules. *Fundam. Appl. Toxicol.* 8:425–31
20. Huff, J. E., Melnick, R. L., Solleveld, H. A., Haseman, J. K., Powers, M., Miller, R. A. 1985. Multiple organ carcinogenicity of 1,3-butadiene in B6C3F1 mice after 60 weeks of inhalation exposure. *Science* 227:548–49
21. Melnick, R. L., Huff, J. E., Chou, B. J., Miller, R. A. 1990. Carcinogenicity of 1,3-butadiene in C57BL/6 X C3H F1 mice at low exposure concentrations. *Cancer Res.* 50:6592–99
22. Matanoski, G. M., Santos-Burgoa, C., Schwartz, L. 1990. Mortality of a cohort of workers in the styrene-butadiene polymer (SBR) manufacturing industry 1943–1982. *Environ. Health Perspect.* 86:107–17
23. Matanoski, G. M., Santos-Burgoa, C., Zeger, S. L., Schwartz, L. 1989. Epidemiologic data related to health effects of 1,3-butadiene. In *Assessment of Inhalation Hazards,* ed. U. Mohr, D. V. Bates, D. L. Dungworth, P. N. Lee, R. O. McClellan, F. J. C. Roe, pp. 201–14. New York: Springer-Verlag. 382 pp.
24. OSHA. 1990. Occupational exposure to 1,3-butadiene; Proposed rule and notice of hearing. Occup. Safety and Health Adm. Dept. Labor. *Fed. Regist.* 55 (155):32736–826
25. Zwickey, R. E., Davis, K. J. 1959. Carcinogenicity screening. In *Appraisal of the Safety of Chemicals in Foods, Drugs and Cosmetics.* Assoc. Food and Drug Off. US. Edit. Comm. Baltimore, MD

26. IARC. 1987. *IARC Monographs on the Evaluation of Carcinogenic Risk to Humans. Overall Evaluations of Carcinogenicity: An Updating of IARC Monographs Vols. 1–42*, Suppl. 7:1–440. Lyon, France: IARC
27. Huff, J. E. 1991. Classification of chemical carcinogens: Levels of evidence of carcinogenicity developed and used by the National Toxicology Program (1983–1990). *Scand J. Work Environ Health* In press
28. Haseman, J. K., Huff, J. E. 1987. Species correlation in long-term carcinogenicity studies. *Cancer Lett.* 37:125–32
29. Huff, J. E., Rall, D. P. 1991. Relevance to humans of carcinogenesis results from laboratory animal toxicology studies. In *Maxcy-Rosenau's Public Health and Preventive Medicine*, ed. J. M. Last, Norwalk, Connecticut: Appleton-Century-Crofts. 13th ed. In press
30. Rao, G. N., Huff, J. E. 1990. Refinement of long-term toxicity and carcinogenesis studies. *Fundam. Appl. Toxicol.* 15:33-43
31. Gould, S. J. 1986. Certainty is unattainable in science. *Science* 7:52
32. Boyland, E. 1969. The correlation of experimental carcinogenesis and cancer in man. *Prog. Exp. Tumor Res.* 11:222–34
33. Maronpot, R. R., Haseman, J. K., Boorman, G. A., Eustis, S. E., Rao, G. N., Huff, J. E. 1987. Liver lesions in B6C3F1 mice: The National Toxicology Program experience and position. *Arch. Toxicol. Suppl.* 10:10–26
34. Newberne, P. M. 1982. Assessment of the hepatocarcinogenic potential of chemicals: Response of the liver. In *Toxicology of the Liver*, ed., G. Plaa, W. R. Hewitt, pp. 243–90. New York: Raven. 338 pp.
35. Anderson, M. W., Maronpot, R. R., Reynolds, S. H. 1988. Role of oncogenes in chemical carcinogenesis: extrapolation from rodents to humans. In *Methods for Detecting DNA Damaging Agents in Humans: Applications in Cancer Epidemiology and Prevention*. ed., H. Bartsch, K. Hemminki, I. K. O'Neill. IARC Sci. Publ. 89:477–85. Lyon, France: IARC. 518 pp.
36. Reynolds, S. H., Stowers, S. J., Patterson, R. M., Maronpot, R. R., Aaronson, S. A., Anderson, M. W. 1987. Activated oncogenes in B6C3F1 mouse-liver tumors: implications for risk assessment. *Science* 237:1309–16
37. Reynolds, S. H., Stowers, S. J., Patterson, R. M., Maronpot, R. R., Anderson, M. W. 1988. Oncogene activation of spontaneous and chemically induced rodent tumors: implications for risk analysis. *Environ. Health Perspect.* 78:175–77
38. Reynolds, S. H., Stowers, S. J., Maronpot, R. R., Anderson, M. W., Aaronson, S. A. 1986. Detection and identification of activated oncogenes in spontaneously occurring benign and malignant hepatocellular tumors of the B6C3F1 mouse. *Proc. Natl. Acad. Sci. USA* 83:33–37
39. Wiseman, R. W., Stowers, S. J., Miller, E. C., Anderson, M. W., Miller, J. A. 1986. Activating mutations in the C-Ha-*ras* protooncogene in chemically induced hepatomas of the male B6C3F1 mouse. *Proc. Natl. Acad. Sci. USA* 83:5825–29
40. Barrett, J. C., Osahimura, M., Koi, M. 1987. Role of oncogenes and tumor suppressor genes in a multistep model of carcinogenesis. In *Critical Molecular Determinants of Carcinogenesis*, ed. H. zur Hausen, J. R. Schleofer. 39:45–56. Austin, Texas: Univ. Austin Press
41. Barrett, J. C., Wiseman, R. W. 1991. Relevance of cellular and molecular mechanisms of multistep carcinogenesis to risk assessment. In *Inferring Carcinogenic Effects in One Species with Data from a Different Species*, ed. D. M. Byrd III, J. D. Wilson. New York: Telford. In press
42. Weinberg, R. A. 1989. Oncogenes, antioncogenes, and the molecular bases of multistep carcinogenesis. *Cancer Res.* 49:3713–21
43. Barbacid, M. 1987. *ras* genes. *Annu. Rev. Biochem.* 56:779–827
44. You, M., Candrian, U., Maronpot, R. R., Stoner, G. D., Anderson, M. W. 1989. Activation of the ki-*ras* protooncogene in spontaneously occurring and chemically induced lung tumors of the strain A mouse. *Proc. Natl. Acad. Sci. USA* 86:3070–3074
45. Barrett, J. C., Wiseman, R. W. 1987. Cellular and molecular mechanisms of multistep carcinogenesis; relevance to carcinogenesis risk assessment. *Environ. Health Perspect.* 76:65–70
46. Barrett, J. C., Huff, J. E. 1991. Cellular and molecular mechanisms of chemically induced renal carcinogenesis. *Renal Failure*. In press
47. Shubik, P., Sicé, J. 1956. Chemical carcinogenesis as a chronic toxicity test. A review. *Cancer Res.* 16:728–42

48. Abelson, P. H. 1987. Municipal waste. *Science* 236:1409
49. Yamasaki, H. 1987. Cancer can be defined simply as a rebellion in an otherwise orderly cellular society. In *Nongenotoxic Mechanisms in Carcinogenesis. Banbury Rep.* 25
50. Eustis, S. L. 1989. The sequential development of cancer: a morphological perspective. *Toxicol. Lett.* 49:267–81
51. Huff, J. E., Eustis, S. L., Haseman, J. K. 1989. Occurrence and relevance of chemically induced benign neoplasms in long-term carcinogenicity studies. *Cancer Metast. Rev.* 8:1–21
52. Off. Sci. Technol. (OSTP). 1985. *Chemical Carcinogens: A Review of the Science and its Associated Principles. Interagency Staff Group on Chemical Carcinogenesis.* OSTP, Exec. Off. President, *Fed. Regist.* 50:10371–442. Also published in *Environ. Health Perspect.* 1986. 67:201–82
53. Doll, R. 1984. Occupational cancer: Problems in interpreting human evidence. *Ann. Occup. Hyg.* 28:291–305