Long-Term Chemical Carcinogenesis Experiments for Identifying Potential Human Cancer Hazards: Collective Database of the National Cancer Institute and National Toxicology Program (1976–1991)

by James Huff* and Joseph Haseman*

The carcinogenicity database used for this paper originated in the late 1960s by the National Cancer Institute (NCI) and since 1978 has been continued and made more comprehensive by the National Toxicology Program (NTP). The extensive files contain, among other sets of information, detailed pathology data on more than 400 long-term (most often 24-month) chemical carcinogenesis studies, comprising nearly 1600 individual experiments having at least 10 million tissue sections that have been evaluated for toxicity and carcinogenicity. Using this data set, we have a) determined the concordance in carcinogenic responses between rats and mice to be 74% and between sexes to be 85% (rats) to 87% (mice); b) discovered that using male rats and female mice would have identified correctly 95% of the positive or no evidence chemical carcinogenicity results obtained using the more extensive protocol; c) established a historical control file of tumor incidence data; d) evaluated the false positive rate in the interpretation of carcinogenesis studies, concluding that this rate is probably no more than 7 to 8%; e) compiled listings of chemicals having like carcinogenic target sites for each of the 37 organs or systems for which histopathology diagnoses have been recorded routinely; f) demonstrated that evaluation of site-specific carcinogenic effects are preferable to doing analyses based on overall (all sites combined) tumor rates; g) learned that few chemicals cause only benign tumors or only liver tumors, the most common target site for chemically induced cancers; h) identified key sources of variability in tumor rates in long-term carcinogenesis studies; i) determined that corn oil gavage or gavage per se exhibits little, if any, adverse impact on long-term studies; j) showed that the Salmonella multistrain assay was as good as a battery of four short-term in vitro tests for predicting in vivo carcinogenicity, yet was only 66% concordant with an 89% positive predictivity and a 55% negative predictivity; k) investigated the relationship between chemically induced toxicity and chemically associated carcinogenicity, finding that few chemicals cause tumors only at the highest exposure concentration. These (and other) derived compilations are most useful for maintaining a historic and objective perspective when evaluating the carcinogenicity of contemporary experiments and for identifying potential carcinogenic hazards to humans.

Introduction

Since the introduction of exposing animals to chemicals in 1918 by Yamagiwa and Ichikawa for detecting chemical carcinogens (1,2), much has been learned about the relevance of these findings for possible effects in humans. Likewise, an enormous amount of knowledge has been gained over the years on how to design, conduct, monitor, evaluate, and interpret the data collected from these carcinogenesis studies (3–7). Evaluating chemicals in laboratory rodents as surrogates for potential human health hazards remains the cornerstone for identifying those chemicals most likely to cause cancer in humans (8–14).

Other than human experience and epidemiological investigations, long-term studies in laboratory animals are the most validated and universally accepted means to determine carcinogenic hazards to the public health and to the environment (e.g., 5,6,8–22). The major public health attribute of these long-term chemical carcinogenesis experiments is to allow better risk assessment (5,9,10) and risk management decisions (23,24) to be made for reducing, preventing, or eliminating exposures to those chemicals identified as constituting real risks to humans (25–27). The findings from these long-term studies assume a major role in the key first step in the risk assessment-risk management process and a contributory part in the second step of the process (9,23,24) (Table 1).

*National Institute of Environmental Health Sciences, P.O. Box 12233, Research Triangle Park, NC 27709.
Address reprint requests to J. E. Huff, National Institute of Environmental Health Sciences, P.O. Box 12233, Research Triangle Park, NC 27709.
Table 1. Risk assessment: comparative definitions.

| National Research Council/National Academy of Sciences, 1983 (23) | Task Force on Health Risk Assessment, 1986 (24) |
|---------------------------------------------------------------|------------------------------------------------|
| Hazard identification: determination of whether a particular chemical is or is not causally linked to particular (adverse) health effects. | Hazard identification: qualitative indication that a substance/condition may adversely affect human health. |
| Dose–response assessment: determination of the relation between the magnitude of exposure and the probability of occurrence of the (adverse) health effects in question. | Hazard characterization: qualitative and quantitative evaluation of the nature of the adverse effects, including their expression as functions of the amount of exposure (dose). |
| Exposure assessment: determination of the extent of human exposure before or after application of regulatory controls. | Exposure characterization: qualitative and quantitative evaluation of the degree of human exposure likely to occur. |
| Risk characterization: description of the nature and often the magnitude of human risk, including attendant uncertainty. | Risk determination: integration of these steps into a scientific determination of the level of risk as a basis for policy consideration. |

As illustrative and important examples of preventive medicine, the International Agency for Research on Cancer (25,28,29) periodically re-evaluates all those chemicals or agents having available human data, and these consensus findings and rationales are published in supplements that contain groupings of chemicals and agents considered to be carcinogenic, probably carcinogenic, or possibly carcinogenic to humans, along with those chemicals not classifiable or probably not carcinogenic to humans (25). Similarly, and as mandated by the U.S. Congress, the Department of Health and Human Services via the U.S. National Toxicology Program (NTP) evaluates the available data and places chemicals into one of two categories: known human carcinogens and those reasonably anticipated to be human carcinogens (27). Once these are identified by either of these organizations, other state, national, and international agencies, as well as public and private interest groups, assess the possible adverse health impact these agents may have on exposed persons. Various actions may be taken to reduce or eliminate exposures to the identified hazard, to locate and study relevant human cohorts for possible carcinogenicity, and to attempt further studies to better characterize the identified potential risk to humans (8,10,12,13).

In the middle 1960s and early 1970s, the National Cancer Institute (NCI) initiated an extensive carcinogenesis bioassay screening program and published in 1976 the first of 200 technical reports (30,31). The NTP continues these efforts, and about 250 more technical reports have been issued or are in press (11,32). These data and various subcombinations form the bases for this overview paper. Parts of this overview come from our previously reported and published analyses and interpretations. For readers wanting more details and perhaps larger bibliographies, please consult the references given at the end of this paper. Most of the papers authored by us or the individual printed technical reports on the chemical carcinogenesis studies are available upon written request.

The overriding aim for doing these individual and collective analyses comes from the never-ending need to more completely understand the fundamental biological principle that extrapolation of experimental and epidemiological findings among the mammalian kingdom is sound science. The ultimate value of these data and resultant interpretations will continue to allow more knowledgeable public, occupational, and environmental health decisions.

Materials and Methods

The chemical carcinogenesis data given in this paper come from the NCI and NTP Technical Reports Series, which describe the results of long-term studies in rodents (11,32). Compared to others who compile information about carcinogenicity of chemicals, our program is unique in that we design, conduct, and evaluate our own experiments, whereas others typically record the results and conclusions made by the various investigators reporting their data in the literature. Further, and most importantly, our data and interpretations are peer-reviewed in public sessions before the technical reports are finalized and made ready for printing and distribution. In total, the database used for this paper comprises 379 long-term chemical carcinogenesis studies involving 1394 individual experiments: male rats, female rats, male mice, and female mice (5,10).

These toxicology studies are typically carried out using both sexes of two species of rodents divided randomly into sets of 50 to 60 animals per group; a control group and two or three exposure concentrations are graduated down from a carefully selected top level (3,5,6,11,19,33–35). The species most often used by the NTP are the inbred Fischer 344 rat and the hybrid B6C3F1 (C3H × C57Bl6) mouse. Duration of exposure is generally 2 years (or about 2/3 the life span of these rodent species). The data, results, evaluations, and interpretations are peer reviewed in public meetings, and this chemical-specific information is made available in the NTP Technical Reports Series (11).

Results and Interpretative Observations

Long-Term Chemical Carcinogenesis Studies

A continuing evaluation of the NCI/NTP database of long-term chemical carcinogenesis studies has resulted in a number of useful scientific findings, several of which are summarized in this section. Nearly 450 chemical-specific studies have been reported, most often involving groups of male and female rats and mice exposed to a chemical for 2 years (11,30–32,36,37). The collated findings are evaluated, interpreted, and presented in public meetings to a nongovernment peer review panel of experts in chemical carcinogenesis.

Each sex-species experimental grouping is given an overall level of evidence of carcinogenicity selected from five categories: two positive levels (clear evidence and some evidence), one uncertain but not negative (equivocal evidence), one for no observed response (no evidence), and one for seriously flawed experiments (inadequate experiment) (38–41).

Roughly one-half of the chemicals induced a positive carcinogenic response in at least one of the several experiments (Table 2). Only 43 of the 313 chemicals (13.7%) studied adequately in male and female rats and in male and female mice caused cancer in all four experimental groups, and 25 (8.0%) produced cancer in one or more sites in three of the four experiments (Table 3). These findings will generally cause a chemical to be placed into the sufficient evidence of carcinogenicity category (25,26) and into the group of substances reasonably anticipated to be carcinogens (27). A prime value of these results permits identification of those chemicals considered...
most likely to be potentially carcinogenic to humans, and thus epidemiological investigations having adequate cohorts could be initiated. As an example, the carcinogenicity of 1,3-butadiene was first determined in animals, and a subsequent epidemiological investigation revealed a positive association between exposure and cancers (8).

### Species Correlations

In these studies, similar results are observed in both species 74% of the time, and the concordance between sexes within a species ranges from 85% (rats) to 87% (mice) (Table 4) (42). To maintain the important contribution of hormonal influence on the carcinogenesis process, we conducted a retrospective evaluation on this question: For how many of the outcomes would we have found the same (a simple qualitative yes [++] or no [- -]) result if we had used a reduced protocol of only one-half the experimental groups? Using male rats and female mice, the answer was that 96% of the chemicals would have been correctly matched to the positive or negative result of the more extended experimental protocol (10). Importantly however, 13 chemicals causing single sex–species positives in either female rats (four) or in male mice (nine) would have been missed, as would 2 chemicals causing carcinogenic effects in both female rats and in male mice (Table 3). For all 15 chemically associated responses, there were only single target organs per sex–species (not considering here equivocal responses), and for only one chemical was the target site sex–specific (uterus).

The major questions that arise from this retrospective evaluation center on whether using this modified protocol in the future on the next group of 400 equally diverse chemicals would a) permit the same percentage of correct responses and b) allow us not to miss any potential carcinogenic hazards to humans. While we contemplate answers to these questions, perhaps one could consider using a modified protocol for certain structural certainties; that is, if an organization had to study a certain nitrosamine, a benzidine or anthraquinone dye, or an aniline compound, for example, then a modified protocol could be possible. For “unknown” chemicals, one would be wise to remain with the historically and currently acceptable paradigm of both sexes of two rodent species.

### Historical Control Tumor Data

Although when evaluating long-term chemical carcinogenesis experiments, the most appropriate control group for comparative and interpretative purposes is always the concurrent control, there are instances in which the use of historical control information can aid an investigator in the overall evaluation of tumor incidence data. One example is for rare (or uncommonly occurring) tumors; another is for a tumor type that shows a marginally significant chemically related increase relative to concurrent controls.

The NTP historical control tumor data file is a “moving window” of studies conducted within approximately a 4 to 5 year time period, and it is updated approximately twice a year. Thus, as new data are entered, the oldest data are removed (but of course not discarded). Periodically, control tumor rates are published for informational purposes (43–45).

The extensive NTP database is an important and frequently used source of information on historical control tumors, and in some respects differs from others in that a) typically the tumor diagnoses and therefore incidence data have been verified by comprehensive quality assurance and pathology review procedures, b) tumor diagnostic nomenclature and criteria across experiments and laboratories are consistent, and c) these data come from one source (albeit several laboratories conduct the studies and record the initial tumor diagnoses), whereas other
available tumor databases often come from myriad locations. In each NTP Technical Report, the relevant historical control tumor incidence data are given for all tumors showing possible chemically related effects. Most important are relatively recent studies carried out at the particular study laboratory. Investigators from industry and other governmental agencies also frequently request use of the historical control database so that they can better interpret and compare the background tumor data of their own studies.

**False Positive and False Negative Rates**

In long-term chemical carcinogenesis studies, approximately 30 to 40 different tissues or organs are examined for possible carcinogenic effects in male and female rats and mice. Under these conditions it is not unusual to find statistically significant changes in tumor incidence that merely reflect random variability. Our database provides a unique opportunity to examine the pattern and frequency of site-specific tumor incidences, and to estimate the likelihood of finding statistical differences by chance alone.

For example, Haseman (46) examined the false positive rate in NTP studies by deriving a statistical decision rule that closely approximated the complex scientific judgment process used in the evaluation of these studies. This "rule" (useful as a guideline, but should not be applied rigidly) is to declare an increase in the incidence in a commonly occurring tumor to be biologically significant if the uppermost exposure level incidence compared to controls is significantly different at the $p < 0.01$ level. For an uncommonly occurring tumor (background rate less than 1 or 2%), a $p < 0.05$ increase is typically required. From this evaluation, the false positive rate associated with this decision rule was demonstrated to be no more than 7 to 8%, implying that the false positive rate in NTP studies is reasonably well controlled. Coupled with biological significance, possible mechanistic relevance, an overall historical perspective, and public peer review, we believe that false positives results have been minimized in our program. A more difficult public health issue that has rarely been considered is the false negative rate within these relatively insensitive assays. Perhaps this will receive more attention in the future.

The NTP database is also quite valuable for the study of statistical decision rules, particularly those that require knowledge of the incidence and variability of historical control data. Although such rules should not be rigidly applied, as noted above, these statistical methods may be useful in that they employ a formal multiple comparisons adjustment to limit the overall false positive rate. For a discussion of statistical decision rules that might be used in the evaluation of laboratory animal carcinogenicity studies, see Haseman (47).

While attention is often focused primarily on controlling false positive rates, one must remember that rodent carcinogenicity studies with only 50 to 60 animals per group are relatively insensitive for detecting weak to moderate carcinogenic responses. Thus, this insensitivity must also be considered when deciding how to guard against false negative outcomes. For example, as mentioned below, we demonstrated that analyses based on overall (all sites combined) rather than site-specific tumor rates would have resulted in a marked increase in the false negative rate. That is, several chemicals showing positive results evaluated using site-specific comparisons would not have been identified using total numbers of tumor-bearing animals (48). Thus, the interpretative emphasis should continue to be on site-specific effects.

**Site-Specific Neoplasia**

Using the NCI/NTP data set we have compiled listings of chemicals having like carcinogenic target sites for each of the 34 organs or systems for which histopathology diagnoses have been recorded routinely (49). The most common tumor site is the liver (15% of all experiments) (50), followed in rank order by lung; hematopoietic system and kidneys; mammary glands; fore-stomach; thyroid glands; Zymbal glands; urinary bladder; and skin and uterus (Table 5) (49). These compilations are most useful for maintaining a historic perspective when evaluating the carcinogenicity of contemporary experiments. Equally important, the chemical–tumor–organ connection permits an evaluation of how well chemically induced cancers in a particular organ in one sex or species will predict or correlate with the other sex or species. Using liver cancers as an example, the overall inter-species concordance is 80%. Likewise, target-site predictions can be made for chemicals selected for study that may be similar to those already evaluated; thereby experimental protocols could be adjusted to allow for example more extensive pathology on preselected target organs (i.e., serial sections of the kidney). Further from these observations, one could decide to use two strains of mice to evaluate a short-chain chlorinated aliphatic compound or to study a human carcinogen in a sex–species known to develop chemically induced tumors in the same site observed in humans. Structural classes of chemicals having a propensity for certain organs can be easily identified from these data. Sex–species responders to particular chemically induced cancers, such as the kidney in male rats, can be recognized at a glance. Like humans, the top 10 sites of cancer are somewhat different between the sexes of rodents. Yet when compared among the mammalian species in general (including humans), the sites and rankings are remarkably similar (10,12,20,49).

**Site-Specific versus Total Tumors**

Using a subset of 81 carcinogenicity studies, we evaluated the issue of interpreting the tumor incidence results based on the proportion of animals with primary tumors (all sites) or the proportion of animals with malignant neoplasms (all sites) compared to analyses of site-specific carcinogenic effects (48). Less than half of the 45 chemicals considered to have induced a carcinogenic response in at least one site in one sex–species experiment showed a significant increase in the incidence of primary tumors combined (22 chemicals) or malignant tumors combined (21 chemicals). Among the 29 chemicals interpreted as not carcinogenic based on site-specific effects, two showed significant increases in overall tumor incidence. Thus, consideration of overall tumor rates seemingly masked several important site-specific effects, while identifying only two additional chemicals apparently missed by using site-specific analyses. These latter two "carcinogenic responses" were not considered to be biologically relevant (48).

In our opinion, at least two major problems are associated with an evaluation based on overall (all sites) tumor rates: combining tumor types of varying morphologies and topographies is
biologically unsound and scientifically questionable, and pooling various tumor types reduces study sensitivity for detecting chemically related increases in site-specific tumor incidences. Further, most national and international guidelines for studying chemicals for carcinogenicity in rodents (or in humans) emphasize site-specific effects (6). Thus, despite purported advantages of analyses based on overall tumor rates (e.g., simplicity; reducing occurrence of false positive and perhaps false negative results), primary emphasis should continue to be placed on site-specific analyses.

**Benign Tumors**

The importance and relevance of benign neoplasia in evaluating the carcinogenic potential of a particular chemical remains an area of active discussion, even though few if any chemicals only cause benign neoplasia. To explore this issue we evaluated the long-term carcinogenesis results for the 143 chemicals reported by the NTP: 81 of these showed neoplastic responses in one or more of the 524 sex – species experiments (51). Of these 81 positive studies, 60 (74 %) were considered positive based on malignant neoplasia; 16 (20 %) were positive due primarily to benign neoplasia, but had supporting evidence of malignant neoplasia in the same organ/tissue; and 5 (6 %) were positive based on benign neoplasia alone. Of the 200 chemicals evaluated by NCI, only 2 were considered to be positive based primarily on benign tumors.

Thus, only 5 of the 143 chemicals evaluated induced benign neoplasia alone (3.5 %), and those observed benign neoplasms are known to progress to malignancy (51, 52). Accordingly, we consider chemically induced benign neoplasia to be an important indicator of a chemical’s carcinogenic potential in rodents, and believe these should continue to be made an integral part of the overall weight-of-the-evidence evaluation process for identifying potential human health hazards.

**Sources of Variability**

A major advantage of an extensive historical control database is that it allows an examination of trends in tumor incidence (and other potentially changeable biological parameters) over time and within and between laboratories, thereby permitting the identification of important sources of variability. This information is useful in the design of long-term chemical carcinogenesis studies and may allow the investigator to address what otherwise be confounding factors in the evaluation and interpretation of the data. Potential sources of variability include the animal room environment, longevity and survival patterns of animals, dietary factors, and differences related to pathology (4, 53).

For example, we learned that the Fischer 344 rats used in our program have shown a steady increase in the background rate of certain commonly occurring neoplasms (e.g., leukemia, and for certain glandular organs such as pituitary, mammary, thyroid, and adrenal glands), together with a decreased survival and an increased body weight (54). A review of slides from early and relatively recent studies identified possible reasons for these shifts. (Every slide from every study is kept in a security archives that allows these historical comparisons and also permits others to view slides to verify our diagnoses; for at least 10 years after a study is reported, the tissues and remaining organs from each animal are retained in case new slides are needed.) The possibilities include a) changes in histopathology diagnostic criteria over time, b) changes in the amount of tissue examined, c) intra- and interlaboratory variability, and d) dietary factors; i.e., the incidence of certain neoplasms (particularly endocrine system and mammary gland tumors) are known to be positively correlated with body weight; differences in protein content influence kidney lesions in the rat.

A similar investigation in B6C3F1 mice (55) did not reveal the same striking time-related changes as seen in rats. The only notable change was an increased incidence of pituitary gland neoplasms in female mice in the more recent studies, which may have been associated with increases in the amount of pituitary gland tissue examined. Additional investigation revealed that liver tumor incidence in mice (particularly in females) was highly correlated with body weight (53). Another factor that might have influenced body weight (thereby perhaps indirectly affecting liver tumor incidence) was a program change in protocol made in 1984 that resulted in mice being housed individual rather than in groups (56).

Our investigation of sources of variability will likely lead to certain other suggested changes in experimental protocol or design. For example, an alternative diet containing (among other modifications) less protein and more fiber is being studied and will probably be adopted in the near future. This should improve survival and reduce tumor rates by limiting the weight gain of the animals, as well as reducing the amount and severity of protein-associated lesions in the rodent kidney. We have also gained a better understanding of the value and limitations of historical control data in the overall assessment of tumor incidence (mentioned earlier).

| Site                | No. positive studies | Site                | No. positive studies | Site                | No. positive studies |
|---------------------|----------------------|---------------------|----------------------|---------------------|----------------------|
| 1. Liver            | 44                   | Liver               | 86                   | Liver               | 104                  |
| 2. Kidney           | 28                   | Lung                | 23                   | Lung                | 30                   |
| 3. Mammary glands   | 22                   | Forestomach         | 17                   | Hematopoietic system| 29                   |
| 4. Zymbal glands    | 18                   | Hematopoietic system| 15                   | Kidney              | 29                   |
| 5. Thyroid glands   | 17                   | Circulatory system  | 11                   | Mammary glands      | 27                   |
| 6. Hematopoietic system | 17         | Thyroid glands      | 9                    | Forestomach         | 23                   |
| 7. Forestomach      | 15                   | Mammary glands      | 9                    | Thyroid glands      | 19                   |
| 8. Skin             | 15                   | Ovary               | 8                    | Zymbal glands       | 18                   |
| 9. Urinary bladder  | 14                   | Harderian glands    | 7                    | Urinary bladder     | 16                   |
| 10. Clitoral glands | 10                   | Uterus              | 7                    | Uterus or skin      | 15                   |
Routes of Exposure

The optimum route of exposure for chemical carcinogenicity experiments should be one that most closely mimics the major human exposure route; this is not always possible, but the attempt should be made. Historically, chemicals have been administered by a single route of exposure: usually in the feed (60% of the time for the data base of the nearly 400 studies we evaluated), but also by oral intubation (27%), by inhalation (5%), by application to the skin (3%), by intraperitoneal injection (3%), or in drinking water (2%)(10). Intubation is generally used for administering unstable, volatile, or reactive chemicals. Microencapsulation is a newly developed technique for administering these agents in feed and will generally be used to replace the gavage technique (57,58). This database permits an evaluation of the possible effects of route of administration on carcinogenic response.

Since oral gavage frequently employs corn oil as the vehicle, a question of particular interest was whether or not corn oil affects tumor incidence. To answer the question, a comparison was made of tumor rates in untreated (or basal diet) controls and corn oil gavage controls, taking into account any interlaboratory variability and time-related trends. Haseman et al. (44) found that corn oil has no apparent effect on tumor incidence in male and female B6C3F1, mice or in female F344 rats. In male F344 rats, however, corn oil was associated with a decreased incidence of mononuclear cell leukemia from approximately 27 to 14% and an increased incidence of acinar cell tumors of the pancreas from approximately 0.2 to 4.4%. These changes appeared related to corn oil rather than to the gavage technique itself because controls from gavage studies using water as the vehicle showed a tumor response for leukemia and pancreatic acinar cell tumors similar to that observed for untreated basal diet controls. Interestingly, acinar cell tumors of the pancreas were not affected uniformly because contemporary controls administered corn oil from the identical batch did not show similar increases. Further, if the two control groups with the largest average body weights are removed from the analysis, the difference for pancreatic tumors between basal and corn oil controls is no longer statistically significant.

A long-term study was designed to evaluate the influence of various levels of corn oil and of select other oils (sesame seed, sunflower seed, and tricaprylin) on the carcinogenic potential in F344 rats. Preliminary results strongly confirm both of the effects noted above. Our increased understanding of the impact (or in most cases, the lack of impact) of corn oil gavage on tumor incidence increases our confidence in the interpretation of studies using corn oil as the vehicle (and gavage as the route of administration). These results may also have implications concerning possible dietary changes to increase the survival of F344 rats (leukemia is the leading cause of death in these animals).

Predicting Carcinogenicity with Genetic Toxicity Assays

Four widely used in vitro genetic toxicity assays were evaluated to determine their individual and collective concordance with the carcinogenicity of a number of chemicals. The end points and short-term assays were mutagenesis in Salmonella typhimurium and in mouse lymphoma cells and chromosome aberrations and sister chromatid exchanges in Chinese hamster ovary cells. The original investigation with 73 chemicals (59) has been followed up with an additional 41 chemicals (60–63).

Results from both studies were similar and are summarized in Table 6 for the combined 114 chemicals. Salmonella performed best among the four short-term tests, achieving a 66% (75/114) concordance, an 89% (32/36) positive predictivity, and a 55% (43/78) negative predictivity of carcinogenicity. Chromosome aberrations also showed a significant association with rodent carcinogenicity, but none was observed for the sister chromatid exchange or mouse lymphoma assays.

Although the performance of individual short-term tests was of interest, the primary focus of the investigation was whether the performance of individual short-term tests for predicting carcinogenicity could be improved by using a battery of two or more tests. Importantly, these data demonstrate, at least for these assays and for this particular 114 chemical data set, that there was no evidence of complementarity among the four short-term tests, and no combination constructed from these assays improved substantially the performance over the Salmonella multistrain assay alone.

These results led our program to adopt Salmonella as the primary in vitro short-term assay to identify mutagens. Use of the three other short-term tests noted above has been curtailed. The NTP is now evaluating in vivo tests for chromosome aberrations and micronuclei as a possible means of predicting or confirming in vitro mutagenicity and in vivo carcinogenicity.

Chemical Toxicity and Carcinogenicity

The possible interrelationships between toxicity (histopathological observations), genotoxicity (i.e., Salmonella), and chemical carcinogenicity in laboratory rodents was investigated using information obtained and reported from 2-year studies involving 99 chemicals (64). Although target organ toxicity and

| Table 6. Pattern of short-term in vitro test results and long-term in vivo carcinogenesis results for 114 chemicals. |
|-------------------------------------------------|
| | Short-term in vitro test and results | Rodent carcinogenicity results |
| | SAL | ABS | SCE | MLA | Positive | Negative |
| 4/4 | + | + | + | + | 22 | 3 |
| 3/4 | + | + | + | - | 0 | 0 |
| | + | + | - | + | 0 | 0 |
| | + | - | + | + | 2 | 0 |
| | - | + | + | + | 4 | 1 |
| | - | + | + | + | 8 | 6 |
| Subtotals | | | | | 22 | 3 |
| 2/4 | + | + | - | - | 2 | 0 |
| | + | - | - | + | 0 | 0 |
| | + | - | + | + | 2 | 0 |
| | - | + | + | - | 2 | 3 |
| | - | + | - | + | 0 | 0 |
| | - | - | + | + | 7 | 12 |
| Subtotals | | | | | 14 | 7 |
| 1/4 | + | + | - | - | 2 | 0 |
| | - | + | - | + | 1 | 1 |
| | - | - | - | + | 3 | 1 |
| | - | - | - | + | 1 | 1 |
| Subtotals | | | | | 9 | 8 |
| 0/4 | - | - | - | - | 2 | 0 |
| | - | - | - | - | 11 | 14 |
| Subtotals | | | | | 11 | 14 |
| Totals | | | | | 67 | 47 |

Abbreviations: SAL, Salmonella typhimurium; ABS, chromosome aberrations (CHO cells); SCE, sister chromatid exchanges (CHO cells); MLA, mouse lymphoma.
carcinogenicity occurred together in some studies (particularly for the nasal cavity and kidney), many chemicals produced carcinogenic effects in organs showing no evidence of toxicity, and conversely, toxicity was frequently observed in organs showing no chemically related carcinogenic effects. The data suggested that only 7 of the 53 chemicals causing cancer in at least one organ of one sex of one species exhibited the types of target organ toxicity that might have been the cause of all the observed carcinogenic effects.

Of the 127 positive sex-species experiments that used multiple doses, 54 (43%) produced statistically significant increases in tumor incidence only at the top dose. However, in 78% (42/54) of these experiments the incidences of site-specific tumors in the lower-dose groups were also numerically elevated relative to controls and were considered to be biologically relevant. Thus, only 9% (12/127) of the carcinogenic effects appeared to reflect a “high-dose only” response. Moreover, no apparent difference in mutagenicity as measured by the Salmonella assay was observed between high-dose only carcinogens and the entire set of chemical carcinogens. In a subsequent correlational analyses of an additional 31 chemicals, the results substantiate those summarized above (65,66).

Although cell proliferation and its possible impact on carcinogenic response were not assessed directly, these results nevertheless suggest that only a relatively small percentage of chemical carcinogens can be identified as possibly acting through an indirect or secondary mechanism. Thus, any attempt to develop generic regulatory policy for chemical carcinogens based solely on the use of 2-year rodent studies to differentiate between “primary” and “secondary” carcinogens will likely be ineffective and/or inaccurate unless specifically backed by relevant mechanistic studies (64).

Melnick (67) evaluated the available information on the influence of cell proliferation (or mitogenesis) on the carcinogenic process in rodents. Although cell proliferation has long been known to have a role in chemically induced tumor development, Melnick (67) concluded that a correspondence between sustained cellular proliferation and carcinogenic response has not been demonstrated adequately to support the hypothesis that chemically induced cell proliferation causes liver cancer. Further, the proliferative response resulting from exposure to many “nongenotoxic” carcinogens is not at all well sustained, yet the elicitation of a carcinogenic response by these chemicals often requires a prolonged exposure duration. Thus, much more research needs to be accomplished before any basic axioms about this important issue can be established (68).

Deductive Conclusions

Long-term chemical carcinogenesis studies must be properly designed, adequately conducted and monitored, subjected to multi-tier pathology diagnoses review, and objectively evaluated and interpreted. In addition to having these essential scientific characteristics, NTP studies undergo a unique peer review process that is conducted in open and public sessions. In our view, this makes the findings and results from these long-term chemical carcinogenesis experiments even more useful and valuable for identifying potential carcinogenic hazards to public health. In addition to alerting us to possibly new chemical concerns and substantiating epidemiological observations, the collective data are valuable and useful for a myriad of other scientific purposes. Some have been given in this paper. Others used or developed by us have been published or are at various stages of being published. In one manner or another most if not all allow us to bring further knowledge to and support for increasing confidence in extrapolating results from animals to humans. The knowledge extends from the clear awareness that all of the chemicals known to induce cancer in humans also cause cancer in adequate long-term animal studies (8,10,12) to the increasing awareness that oncogene activation patterns in human and animal tumors are strikingly similar and represent evolutionary conservation (e.g., 69–71).

In our opinion, the data from these studies together with our experience over the years permit us to make the following conclusions (not given in any particular order):

1. Carcinogenicity findings from experiments in laboratory animals are logical and scientifically reasonable for identifying and predicting potential carcinogenic effects to humans.
2. Most chemicals are not carcinogenic.
3. Chemicals considered to be carcinogens may differ significantly with regard to the strength of the induced carcinogenic response, and for practical purposes can be divided into qualitative groups based on empirical indicators of potency, i.e., exposure concentrations, number of positive experiments, multiplicity of tumor sites, numbers of tumors per site, common versus uncommon tumors, site correspondence across sexes and species, magnitude of incidence rates, dose response patterns, latency, metastases, preneoplastic lesions (see Table 3 for examples).
4. Malignant and benign tumors of the same site and/or type should be combined for interpretation. However, combining biologically unrelated tumors for statistical evaluation should be avoided.
5. Benign tumors induced by chemicals are relevant for judging carcinogenicity; few chemicals produce only benign tumors.
6. Site-specific tumor analyses should be used for determining carcinogenic effects.
7. Liver tumors are appropriate and valid for identifying chemical carcinogens and for determining potential cancer hazards to humans.
8. Cellular toxicity and resultant sequelae are not uniformly associated with chemically induced carcinogenicity. Many chemicals produce carcinogenic responses in organs showing no evidence of toxicity, and, conversely, toxicity frequently occurs in organs exhibiting no chemically related carcinogenic effects.
9. Most chemical carcinogens do not cause cancer only at the highest exposure used.
10. Mechanisms of action are not yet considered to be understood well enough to become significant factors in the evaluations of chemically induced carcinogenesis, but certainly all relevant biologic information (oncogene activation and tumor suppressor genes, pharmacology and pharmacokinetics, hormonal influences, and DNA damage and repair, among others) should be considered when deciding a single level of evidence of carcinogenicity.

II. Cell replication (proliferation, mitogenesis) is an important factor in chemical carcinogenesis, yet scientific data do not sustain the hypothesis that enhanced cellular proliferation
in a particular organ associates consistently with an increased induction of neoplasia.

12. The route of exposure has little or no effect on the inherent carcinogenicity of a chemical, although the specific sites of carcinogenicity may vary with the route of administration used.

13. Corn oil has little or no influence on whether or not a chemical is or is not a carcinogen.

14. Interspecies concordance in carcinogenic response is good, yet at least two species should in most cases continue to be used for identifying chemical carcinogens.

15. Salmonella mutagenicity has a relatively low concordance with rodent carcinogenicity.

16. Cage location and related factors have no important impact on chemical carcinogenicity.

17. Chemical structure in general does not allow for cancer prediction in rodents or humans.

18. False positive rates associated with evaluating chemically induced carcinogenic effects in rodents are low. Unfortunately, little is known about the false negative rate, which is of more importance to public health.

In conclusion, much has been learned from these largely observational studies that permits us to make the above statements, all of which are based on the data and facts garnered from the extensive chemical carcinogenesis database developed in sequence by the National Cancer Institute and by the National Toxicology Program. This database is unique in the sense that not only is it the single largest available collection of program-generated chemical carcinogenesis facts and figures, but it is perhaps the only system containing carcinogenesis (and other) data obtained from experiments that have been conducted and evaluated using a reasonably consistent core of design protocols, histopathological diagnostic criteria, and interpretational guidelines (all of which have been established and honed by us over the years to keep reasonable pace with new and relevant scientific advancements).

Another unique character of our data collection and complementary pathology archives of tissue specimens and histopathology slides is that we make the original data and slides available to others to use in any way they wish. We also invite individuals and organizations to study and scrutinize the raw data and tissue specimens and sections that represent the actual experimentally generated information sources of the database.

Staff consistency and balanced scientific standards, together with a public peer review process and a keen awareness of the public health ethos must also be considered unique, desirable, and supportive of the value of this particular data collection.

We thank and dedicate this paper to David Rall, who has recently retired from the U.S. Public Health Service after an exemplary career as a physician and scientist. We appreciate his leadership, his public health and environmentally oriented conscience, and his belief and practice that good research comes from good scientists who not only must be supported but who must also be allowed to pursue their research efforts unencumbered and with resolute freedom.

We also thank Jef French and Chris Portier for reviewing this manuscript and for offering valuable suggestions. Donna Mayer was most helpful throughout the preparation of this paper.

REFERENCES

1. Yamagiwa, K., and Ichikawa, K. Experimental study of the pathogenesis of carcinoma. J. Cancer Res. 3: 1–29 (1918).

2. Shimkin, M. B. Some Classics of Experimental Oncology: 50 Selections 1775–1965. NIH Publication No. 80-2150. U.S. Department of Health and Human Services, Washington, DC, 1980.

3. Chhabra, R. S., Huff, J. E., Schwartz, B. S., and Selkirk, J. An overview of pre-chronic and chronic toxicity/carcinogenicity experimental study designs and criteria used by the National Toxicology Program. Environ. Health Perspect. 86: 313–321 (1990).

4. Haseman, J. K., Huff, J. E., Rao, G. N., and Eustis, S. L. Sources of variability in rodent carcinogenicity studies. Fundam. Appl. Toxicol. 12: 793–804 (1989).

5. Huff, J. E. Design strategies, results and evaluations of long-term chemical carcinogenesis studies. Scand. J. Work Environ. Health, 18(Suppl. 1): in press.

6. Montesano, R., Bartsch, H., Vainio, H., Wilbourn, J., and Yamasaki, H. Long-Term and Short-Term Assays for Carcinogens. A Critical Appraisal. IARC Scientific Publications No. 83. International Agency for Research on Cancer, Lyon, 1986.

7. Rao, G. N., and Huff, J. E. Refinement of long-term toxicity and carcinogenicity studies. Fundam. Appl. Toxicol. 15: 33–43 (1990).

8. Huff, J. E., and Rall, D. P. Relevance to humans of carcinogenesis results from laboratory animal toxicology studies. In: Marcy–Rosenau–Last’s Public Health and Preventive Medicine, 13th ed. (J. M. Last and R. B. Wallace, Eds.), Appleton-Century-Crofts, Norwalk, CT, 1992, pp. 433–440, 453–457.

9. Huff, J. E., and Noel, D. G. Hazard identification. Perspective and overview on the concepts and value of the initial phase in the risk assessment process of cancer and human health. Scand. J. Work Environ. Health, 18(Suppl. 1): in press.

10. Huff, J. E., Haseman, J. K., and Rall, D. P. Scientific concepts, value, and significance of chemical carcinogenesis studies. Ann. Rev. Pharmacol. Toxicol. 31: 621–652 (1991).

11. Huff, J. E., McConnell, E. E., Haseman, J. K., Boorman, G. A., Eustis, S. L., Schwartz, B. A., Rao, G. N., Jameson, C. W., Hart, L. G., and Rall, D. P. Carcinogenesis studies: results from 398 experiments on 104 chemicals from the U.S. National Toxicology Program. Ann. N.Y. Acad. Sci. 534: 1–30 (1988).

12. Tomatis, L., Aitio, A., Wilbourn, J., and Shuker, L. Human carcinogens so far identified. Jpn. J. Cancer Res. 80: 795–807 (1989).

13. Tomatis, L., Aitio, A., Day, N. E., Heselteine, E., Kaldor, J., Miller, A. B., Parkin, D. M., and Riboli, E. Cancer: Causes, Occurrence and Control. IARC Scientific Publications No. 100. International Agency for Research on Cancer, Lyon, 1990.

14. Office of Science and Technology, Interagency Staff Group on Chemical Carcinogenesis. Chemical carcinogens: a review of the science and its associated principles. Fed. Reg. 50: 10371–10442; Environ. Health Perspect. 67: 201–232 (1986).

15. NRC/NAS. Chemical contaminants: safety and risk assessment. In: Drinking Water and Health. National Research Council/National Academy of Sciences, National Academy of Sciences, Washington, DC, 1977, pp. 19–62.

16. Occupational Safety and Health Administration. Identification, classification and regulation of potential occupational carcinogens. Fed. Reg. 45: 5001–5296 (1980).

17. OTA, Congress of the United States. Assessment of technologies for determining cancer risks from the environment. Office of Technology Assessment, Washington, DC, 1981.

18. OTA, Congress of the United States. Identifying and Regulating Carcinogens. Office of Technology Assessment, Washington, DC, 1987.

19. NTP Ad Hoc Panel. Report on an Ad Hoc Panel on Chemical Carcinogenesis Testing and Evaluation. Presented to the NTP’s Board of Scientific Counselors, National Toxicology Program, Research Triangle Park, NC, 1984.

20. Muir, C. Epidemiology, basic science, and the prevention of cancer: implications for the future. Cancer Res. 50: 6441–6448 (1990).

21. Hakama, M., Beral, V., Cullen, S. W., and Parkin, D. M., Eds. Evaluating Effectiveness of Primary Prevention of Cancer. IARC Scientific Publications No. 103. International Agency for Research on Cancer, Lyon, 1990.

22. Huff, J. E. Applicability to humans of rodent-specific sites of chemical carcinogenicity: tumors of the forestomach and of the Harderian, preputial and Zymbal glands induced by benzene. J. Occup. Med. Toxicol. 1: 97–130 (1992).

23. NRC/NAS. Risk Assessment in the Federal Government: Managing the Process. National Research Council/National Academy of Sciences, National Academy Press, Washington, DC, 1983.
