Human Carcinogenic Risk Evaluation: An Alternative Approach to the Two-Year Rodent Bioassay

Samuel M. Cohen

Department of Pathology and Microbiology, University of Nebraska Medical Center, 983135 Nebraska Medical Center, Omaha, Nebraska 68198-3135

Received March 29, 2004; accepted April 21, 2004

An approach to the evaluation of carcinogenic risk resulting from exposure to a given chemical is presented in place of a reliance on two-year rodent bioassays. An emphasis is placed on evaluation of the potential DNA reactivity or increased cell proliferation that can be produced by a chemical. The special cases of immunosuppressive and estrogenic chemicals are considered. These evaluations are proposed to involve a combination of in vitro assays, computerized models, and short-term (up to 13 weeks) bioassays in rodents. The emphasis is on mechanistic understanding and evaluation of the dose response and relevance to humans.

Key Words: mechanism; DNA reactive; non-genotoxic; estrogen; immunosuppressive; cell proliferation.

Chemicals have been known to cause human cancer for more than two centuries. For most of that time we have identified specific chemicals and mixtures as carcinogens based on clinical observations and epidemiologic investigations. During the past half century, however, the focus has shifted from identifying carcinogens in the environment to preventing exposure to them in the first place. For this, we have relied increasingly on a variety of hazard identification screening models, along with an increased understanding of mechanisms of action by which these chemicals can cause cancer.

In the 1950s and 1960s, several chemicals were identified as carcinogens in rodents that were also known to be carcinogenic in humans. The two-year bioassay was developed to provide a standardized screening procedure for evaluating chemicals with the assumption that these were predictive of human carcinogenic risk. It is important in interpreting the results from any of these models to keep in mind the dictum of the famous mathematician, George Box (1979), “All models are wrong. Some are useful.” As scientists, it is our responsibility to ascertain what is the useful information from these models and what information cannot be extrapolated to humans.

For carcinogenicity risk assessment, it is important to keep focused on the ultimate aim of these investigations: Does the agent cause cancer in humans?

In utilizing animals as a bioassay screening model, two fundamental assumptions are made: (1) the results we observe in the animal model are relevant to humans (species extrapolation); and (2) the dose administered to the animals is relevant to the exposure levels in humans (dose extrapolation). For many chemicals, particularly DNA reactive carcinogens, these assumptions are reasonable. However, as has been clear for many chemicals, one or both of these assumptions are not appropriate.

We are aware of the difficulties of the two-year bioassay, including the extraordinary cost, now totaling several million dollars for a properly conducted good laboratory practices compliant bioassay in rats and mice. To do a proper two-year bioassay, including preliminary dose setting studies and final evaluations, these studies can take four to five years to complete. The issue of the extraordinarily high doses (maximum tolerated dose) usually used in these bioassays has been the topic of considerable debate. Most importantly, however, is the increasing suspension that the findings in these rodent bioassays are not relevant to human risk.

Several alternatives to the two-year bioassay have been proposed, some involving animal models whereas others have involved in vitro or computerized modeling. Numerous genotoxicity assays have been developed as screens for carcinogenicity, beginning with the Salmonella mutagenicity assay. However, it has become clear that such assays predict carcinogenicity relatively poorly, and only for a subset of carcinogens, namely DNA-reactive chemicals. Computerized chemistry structure activity relationship models have been developed, increasingly useful since an extraordinary amount of data has been accumulated to form the foundation for these models. However, these models tend to be restricted to specific end points, such as DNA reactivity, mutagenesis, receptor interactions, and others. Predictability for carcinogenicity is generally not very strong.

New animal bioassays for screening hazard identification have been developed, based either on mechanistic approaches to animal bioassays, such as the Ito medium-term bioassay in rats, or utilizing transgenic or knockout mice that incorporate or eliminate a gene known to be related to human cancer (e.g., as for the heterozygous p53 mouse). However, these models require careful interpretation since there remains the issue of species
differences in toxicokinetics, metabolism, anatomy, and physiology. No matter how “humanized” some of these models are designed to be, they are still rodents. That is not to say that such models are without value, particularly in exploring mechanisms of action.

DNA REACTIVITY AND CELL PROLIFERATION

Several years ago, we developed a model of carcinogenesis based on first principles known about carcinogenesis (Greenfield et al., 1984). Our assumptions were that genetic alterations are required for cancer formation, more than one such alteration is essential, and DNA replication does not have 100% fidelity. DNA replication is extraordinarily precise largely because of extensive repair mechanisms that have evolved. Nevertheless, mistakes occur. Spontaneous errors in humans have been estimated at approximately one per 10^{10} nucleotides per DNA replication, which translates to less than one mistake per DNA replication for the human genome. During the past few decades the basis for DNA replication errors have been identified, including oxidative damage, depurination and depyrimidination, deamination, nitric oxide interaction, endogenous exocyclic adducts, and others. Although these events occur many hundreds to thousands of times daily in each cell, most are repaired, but not all.

Based on these fundamental assumptions, there are only two basic means by which an increase in carcinogenic risk can occur: (1) increase the net rate of DNA damage per cell division (DNA reactivity) or (2) increase the number of DNA replications, increasing the likelihood of “spontaneous” replication errors. Numerous pathways have been identified for metabolic activation of chemicals to reactive intermediates that combine to form DNA adducts, some of which are mutagenic. If the adduct forms in a critical part of a gene that is essential to development of cancer, the probability of carcinogenesis increases.

Likewise, there are numerous pathways by which the number of DNA replications can be increased (Cohen, 1998). The essential parameter is the number of DNA replications, not necessarily the rate of DNA replication. The number of DNA replications can be affected by either an increase in the rate or an increase in the total number of cells in the target population.

An increase in cell proliferation can occur by either increasing cell births or decreasing cell deaths, the latter leading to an increased accumulation of cells. Increased cell births can occur either by direct mitogenesis, associated with hormonal or growth factor stimuli, or by toxicity and ensuing regeneration. Decreasing cell deaths can occur by either inhibiting apoptosis (a normal event in many tissues) or inhibiting cell differentiation (fundamentally a cell death process).

The DNA damage has to occur in the pluripotential cell population of the tissue, not cells that are committed to differentiation or fully matured. For many tissues, the specific cell population that represents the target for tissue repair and for carcinogenesis has not been precisely identified. The distinction between proliferation of the pluripotential cell type versus a maturing cell type is best illustrated in comparison of hyperplastic and adenomatous polyps of the colon. Adenomatous polyps are precancerous lesions composed of a proliferation of crypt cells, the colonic pluripotential cell population. In contrast, hyperplastic polyps are not preneoplastic lesions, and are composed of proliferating cells maturing or fully differentiated cells of the colonic gland.

HUMAN RELEVANCE

Regardless of whether a chemical increases carcinogenic risk by directly damaging DNA or increasing cell proliferation (or both), the results of animal bioassays need to be rigorously evaluated for relevance to human cancer risk. Frameworks for such an evaluation have been proposed, including a recent one developed by a working group organized by the International Life Sciences Institute/Risk Science Institute sponsored by the U.S. Environmental Protection Agency (EPA) and Health Canada (see the November 2003 issue of Critical Reviews in Toxicology). This human relevance framework consists of three fundamental questions:

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

Once an analysis has been performed, a statement regarding confidence in the determinations and their implications is made. If a chemical induces more than one type of tumor, the MOA for each tumor type is evaluated separately, even if it ends up that they have a common MOA.

It is important in formulating the MOA that specific key events are identified and evaluated. A framework for evaluation of an animal MOA was established by the International Programme on Chemical Safety (IPCS) and the U.S. EPA (Sonich-Mullin et al., 2001). This framework includes a brief description of the animal MOA based on a clear description of each of the key events that underlie the MOA. Dose response relationships, temporal associations, strength, consistency, and specificity of association of key events in tumor response, biological plausibility and coherence, and overall conclusions about the MOA are evaluated. Alternative modes of action are also thoroughly evaluated. A major consequence of such an evaluation of MOA is that it clearly can identify remaining uncertainties and data gaps requiring further exploration. For many chemicals, an adequate description of the key events of the MOA has been developed.

To evaluate the second question regarding plausibility of the key events in humans, a concordance table was found to be a useful tool. The key events are listed sequentially, evidence for occurrence of key events in the animal are tabulated, and finally, an evaluation of whether the key events can occur in humans. If any one of the key events cannot occur in humans, the MOA is not qualitatively relevant to the human situation.
Several examples have been identified for which the animal MOA is not relevant to humans. Most well known is the example of $\alpha_2\mu$-globulin-related induction of male rat kidney tumors, based on interaction of the administered chemical (or metabolite) with $\alpha_2\mu$-globulin, ingestion, impaired digestion and accumulation in the renal tubular cells leading to cell death, regeneration, and eventually renal cell tumors. No comparable protein is present in humans, and therefore, this MOA cannot occur in humans. Similarly, formation of calcium phosphate-containing urinary precipitate following administration of high doses of sodium salts to rats, such as saccharin or ascorbate, occurs by a mechanism that does not occur in humans. This lack of human relevance led to the down classification of saccharin by the International Agency for Research on Cancer (IARC) and delisting from the National Toxicology Program List of Carcinogens.

The third question evaluates quantitative aspects of the MOA, including kinetic and dynamic factors. Again, a side-by-side concordance table is a useful approach.

An example of quantitative differences is the induction of thyroid tumors in rats by agents, such as phenobarbital, that induce metabolism of thyroid hormones, leading to a feedback elevation of the thyroid stimulating hormone (TSH) levels that stimulate rat thyroid follicular cell proliferation. Although this sequence of events can theoretically occur in humans, there is no evidence quantitatively that there is sufficient induction in humans to produce a change in TSH levels, and thyroid hormone physiology differs greatly in rodents and humans. Also, increased TSH levels produced by other mechanisms do not lead to follicular cell proliferation in humans, but rather, there is hypothyroidism of the follicles instead. Thus, an evaluation of the kinetic and dynamic aspects of the process in humans leads to the conclusion that this MOA is not relevant to humans.

An evaluation of the level of confidence in the concordance analysis between animals and humans is required. Like the animal MOA framework, such an analysis can identify uncertainties and data gaps in the evaluation of human relevance, which can then be investigated further. It has become common for laboratories, whether government, academic, or in industry, to explore the MOAs of chemicals found to be positive in two-year rodent bioassays to evaluate potential human relevance.

Even if the overall MOA is qualitatively and quantitatively relevant to humans, an evaluation of the dose response may indicate that expected human exposure levels do not represent a cancer risk.

Formation of urinary calculi is a well-known mechanism in rodents for producing bladder cancer. The key events include ingestion of adequate amounts of the chemical leading to sufficient levels in the urine to produce precipitation, formation of calculi, urothelial toxicity, regeneration, and ultimately tumors. A similar mechanism can occur in humans, and there is epidemiologic evidence supporting a slight increased relative risk in humans exposed to urinary tract calculi. However, for a variety of anatomic, physiologic, and dynamic reasons, the risk in humans is significantly less than in rodents. Most importantly, this is a high-dose phenomenon only. Thus, even though the MOA is relevant to humans, human cancer risk is dependent on exposure levels. For example, melamine produces calculi and bladder tumors in rats. However, exposure in humans is four to five orders of magnitude lower than what is necessary for the formation of urinary tract calculi. Thus, at present usage levels of melamine there is not a carcinogenic risk to humans. A distinction exists between evaluation of the MOA in hazard identification and relevance to risk assessment.

It is quite apparent in evaluating the MOAs of many chemicals and applying the framework for human relevance that the carcinogenic process involves several preneoplastic key events which can be identified and investigated in significantly shorter term experiments than waiting for the ultimate development of cancer. Rather than spending the time and financial resources performing the two-year bioassays, I am proposing that the focus be on an evaluation of preneoplastic events, determination of their human relevance, and evaluation of the dose response compared to probable human exposures.

**RODENT SPECIFICS**

In extrapolating from rodents to humans, several variables have been identified which are essential for the extrapolation.

Some organs occur in rodents for which there is no human counterpart. These include Zymbal’s gland, the Harderian gland, and the forestomach. For DNA reactive substances, formation of tumors in these organs may be indicative of a more generic carcinogenic hazard. However, for non-DNA reactive substances, formation of tumors in these rodent tissues do not have relevance to humans.

Some argument could be made regarding the forestomach as representing the esophagus in humans. However, since the rodent forestomach has a morphologically different mucosa than the human esophagus, the forestomach has a keratinized squamous epithelium whereas the esophageal mucosa is nonkeratinized. In addition, the forestomach mucosa is exposed to high levels of acidity normally for which defense mechanisms have been established so that the tissue is not damaged. In contrast, acid is a known toxicant to the human esophageal mucosa, commonly seen in individuals with gastroesophageal reflux. This line of reasoning was used in the evaluation of the carcinogenicity of butylated hydroxyanisole (BHA), which produces forestomach cancer in several rodent species. These tumors were concluded not to be relevant to humans, and BHA remains as a commonly used food preservative (Grice, 1988).

In addition, certain tumors occur in rodents for which there is not a human equivalent, including splenic mononuclear cell leukemia in rats (particularly F344 rats) and the mouse submucosal mesenchymal lesion of the urinary bladder (also seen in seminal vesicles and uterus). Since these lesions do not have a human counterpart, their induction by a chemical, particularly if...
non-DNA reactive, in rodents is not relevant to human cancer risk.

Rodent endocrine tumors also appear to have little or no human relevance even though feedback mechanisms are qualitatively similar. However, specific control mechanisms and other endocrinologic and physiologic factors are so different, especially quantitatively, that most have little relevance to humans. Tumors of the thyroid, adrenal cortex and medulla, anterior and posterior pituitary, parathyroid, pancreatic islets, and gastrointestinal endocrine cells are frequently targets of chemical carcinogens in rodents, but have uniformly been demonstrated not to have relevance to humans, at least for non-DNA reactive substances.

Similarly, endocrine tumors occurring in reproductive organs, such as granulosus cell tumors of the ovary, Leydig cell tumors of the testes, and tumors of the rodent prostate or breast have considerably different hormonal and physiologic controls than in humans. Results with non-DNA reactive carcinogens towards these tissues in rodents are generally not applicable to humans.

In contrast, for the liver numerous DNA reactive and non-DNA reactive MOAs have been identified, including DNA reactivity, peroxisome proliferation, $P450$ induction, metal overload and oxidative damage, or cytotoxicity and regeneration, some of which appear to be relevant to humans and some do not. These modes of action can readily be identified in short-term evaluations ranging from a few days to a few months.

Similarly, several MOAs for the kidney and lower urinary tract have been identified, some of which are relevant to humans and some are not. In the kidney, MOAs include DNA reactivity, cytotoxicity and regeneration, increased apoptosis and regeneration, $\alpha_2\varepsilon$-globulin binding, enhanced aging nephropathy, and metal overload with consequent oxidative damage. For the lower urinary tract, MOAs include DNA reactivity, direct or indirect formation of urinary solids, alterations in urinary composition, cytotoxicity and regeneration, or direct mitogenesis. Again, some are relevant to humans and some are not.

ASSESSMENT OF HUMAN CARCINOGENIC RISK

Taking all of the above information into consideration, I am postulating that an assessment of the potential human carcinogenicity for a chemical or mixture can be evaluated without performing a two-year rodent bioassay.

To begin with, an evaluation of the genotoxicity of a chemical is required. However, I believe that it is essential to focus on the potential for DNA reactivity rather than the broader implications involved with the term genotoxicity. DNA reactivity has significantly different implications with regard to dose response and extrapolation to humans. Many of the other mechanisms of genotoxicity, such as micronucleus formation, chromosomal aberrations, clastogenicity, effects on DNA repair, or effects on the mitotic apparatus, involve an indirect effect on DNA and may or may not be relevant to human carcinogenesis. Importantly, these indirect mechanisms clearly imply a nonlinear dose response, and in many instances, imply a threshold response.

To assess DNA reactivity, one can utilize the Ames assay. Also, excellent computerized structure activity relationship models are being established, the latter being particularly useful for evaluation of metabolites of a chemical. An in vivo evaluation can be made by specifically investigating formation of DNA adducts. This is a more difficult task, but can be addressed in many instances. If DNA reactivity is identified, an evaluation of whether the metabolic process in rodents is relevant to humans can be addressed along with the dose response and possible nonlinearities, saturation of enzyme kinetics or other influences that might affect a quantitative assessment in humans.

The next question to be addressed is whether the chemical increases cell proliferation. There are two special circumstances that can be addressed before a general evaluation of cell proliferators: immunosuppression and estrogenic activity.

Immunosuppressive chemicals are known to be carcinogenic in humans. However, unlike what was originally hypothesized for the immune surveillance theory, it is not surveillance of neoplasia (tumor specific antigens) by which this process acts (Cohen et al., 1991). Rather, immunosuppression at clinically relevant levels produces decreased surveillance of infectious organisms and consequent disease processes. Some of these infections or reactivation of previous infectious organisms are related to the development of tumors, such as Epstein-Barr virus (EBV) and B-cell lymphomas, herpes virus 8 related to Kaposi’s sarcoma, human papilloma virus (HPV) related to squamous cell carcinoma (especially of the cervix), human T-cell leukemia viruses (HTLV), and hepatitis B and C viruses (HBV and HCV) related to development of hepatocellular carcinomas. Regardless of the mechanism of immunosuppression, whether by exogenous chemical, inherited disorder, or development of acquired immunodeficiency syndrome (AIDS), the individual has an increased susceptibility to developing cancer. It is not an increased susceptibility to all cancers, only those related to specific infectious organisms. Thus, if a chemical is non-DNA reactive but is known to be immunosuppressive to the level of clinical relevance, we can assume that it will be carcinogenic in humans without having to further evaluate the chemical in rodent bioassays. However, short-term evaluations in rodents or other species can be utilized for identifying other types of toxicity.

Similarly, there is considerable information to indicate that clinically significant levels of estrogenic activity are related to an increased incidence of certain tumors in humans, especially endometrial and breast carcinomas. If a chemical is identified as having significant estrogenic activity, one can assume that it will be potentially carcinogenic in humans. Short-term assays can be used to determine dose response and other specifics for predicting human risk, including estrogenic activity relative to human exposure levels.
Increased cellular proliferation beyond these two examples can be evaluated in a four and/or thirteen week bioassay in rodents, utilizing already available techniques. However, this requires considerably more detailed evaluation than is currently used for such assays, and would include detailed evaluation of preneoplastic, proliferative or toxic lesions, organ weights, blood and urine chemistries, and bromodeoxyuridine (BrdU) labeling indices (probably utilizing 1 h and three day administrations to evaluate different tissues). Specialized studies, such as assays for evaluation of aberrant crypt foci, immunohistochemistry, or others could be included as appropriate. It is likely that future developments in genomics will provide useful tools to further evaluate these issues, but will also likely raise difficulties of interpreting nonadverse effect versus adverse effect. Importantly, if there is a signal in any of these short-term studies, it can be further evaluated in considerable detail, including dose response, metabolism and kinetics, and an evaluation of the human relevance.

Some might argue that such analyses require considerable cost and time. I can only remind such individuals that the current two-year bioassay requires millions of dollars and four to five years to complete. Once the results of the two-year bioassay become available, we have to perform the analysis that I am indicating for evaluation of the human relevance of the MOA, anyway.

A summary of what I am proposing is presented in Figure 1. Based on technologies already available and on our understanding of carcinogenesis, particularly on our knowledge and identification of preneoplastic changes, I strongly believe that it time for us to stop performing the two-year rodent bioassay as a routine screening procedure for cancer hazard identification of chemicals.

**REFERENCES**


**FIG. 1.** A proposed guide for evaluating the potential carcinogenicity of chemicals. Each box poses an evaluation to be performed. If the sequence results ultimately in a No that is in a circle, there is no (or negligible) carcinogenic risk in humans. If the sequence results ultimately in a Yes that is in a triangle, it poses a presumptive human carcinogenic risk.