

Historical perspective of the cancer bioassay

by Timothy Pastoor, PhD,¹ James Stevens, PhD²

Pastoor T, Stevens J. Historical perspective of the cancer bioassay. *Scand J Work Environ Health* 2005;31 suppl 1:129–140.

For more than half a century the rodent bioassay has been used in an attempt to predict the potential of chemicals to cause cancer in humans. High mortality and animal husbandry problems led to a more statistically correct study that utilized 50 or more animals per group, cost 1 to 3 million US dollars, and required 3 to 6 years to complete. With an increased understanding of the carcinogenic process came the realization that chemically induced cancer in rodents may not be predictive of carcinogenic potential in humans. This paper describes the odyssey of the bioassay and recommends design improvements that include the incorporation of new molecular methods, the development of data to describe a plausible mode of action, pharmacokinetic and dynamic data for cross-species extrapolation, relevant dose levels consistent with human exposure scenarios, reduction in the number of animals used, and a harmonized international testing strategy for the acquisition of data useful in risk assessments.

Key terms concordance; harmonization; plausibility; predictive; refinement; review.

Although epidemiologic studies had their beginnings as early as 1775, when Dr Percivall Pott, a surgeon at St Bartholomew's Hospital in England, ascribed scrotal cancer to contamination of the skin by chimney soot (1), it was not until 1915 when Yamagiwa & Ichikawa (2) produced malignant epithelial tumors through the application of coal tar to the ears of rabbits that cancer studies with animals began. As it was recognized that coal tar was a human carcinogen, this critical experiment led to other known human carcinogens being administered to animals to elicit tumors to validate the approach. This practice evolved into the use of animal surrogates for evaluating the potential for chemicals to cause cancers in humans.

The usefulness of animal tests is usually justified by reference to the fact that all compounds identified as human carcinogens with the use of epidemiologic methods have also been demonstrated to be animal carcinogens (3–5). Whereas the connection between positive epidemiologic findings and positive findings in animals has shown consistency, the reverse correlation does not necessarily hold. Positive findings in animals typically provide only an alert to the potential of chemically induced cancer in humans. Thus the use of the bioassay has shifted from the verification of human cancer causation to one of prediction. As such, the bioassay has

become a centerpiece in human health risk assessment and the regulation of chemicals.

Animal bioassays have been used for over three quarters of a century and have been used by regulators for 40 years for human hazard assessment and about half of that time for quantitative risk assessment. In the United States, the Food and Drug Administration (FDA), the National Cancer Institute, the National Toxicology Program (NTP), and the Environmental Protection Agency (EPA) have been the most significant players in the area of human health and cancer risk assessment. In Europe, the World Health Organization's International Agency for Research on Cancer (WHO/IARC) has been the prime mover.

Given the predictive role of the bioassay in risk estimation, several assumptions set the foundation for the design and interpretation of the bioassay. At least six primary assumptions are (i) that the carcinogenic response observed in animals can usually be directly translated to humans and tissue site concordance is not necessary (6), (ii) that the relationship between dose and cancer response is linear, (iii) that the slope of the dose–response relationship at low doses can be derived from data at high doses, (iv) that large doses, including the maximum tolerated dose (MTD), must be employed to insure reasonable power for tumor detection (7), (v) that a sufficient number of animals (50 rodents per gender

¹ Syngenta Crop Protection Inc, Greensboro, North-Carolina, United States.

² Wake Forest School of Medicine, Winston-Salem, North-Carolina, United States.

Correspondence to: Dr Timothy Pastoor, Syngenta Crop Protection Inc, Greensboro, NC, 27419, USA. [E-mail: tim.pastoor@syngenta.com]

and group) should be included in the experimental design to insure reasonable power for detecting carcinogenic effects (7); and (vi) that all carcinogens can be treated in the same manner, regardless of the mode of action.

Bioassay design

These assumptions have dictated the experimental design and have also rigidly defined the boundaries of interpretation. The standard rodent cancer bioassay conducted by the NTP of the National Cancer Institute generally included at least 50 animals per gender and group and at least four groups (control, 1/4 the MTD, 1/2 the MTD, the MTD) or a minimum of 400 animals (7, 8). For each material tested, bioassays were carried out in the strains of choice of the NTP, B₆C₃F₁ mice and Fischer 344 rats (9). A typical experimental design for an NTP study is presented in table 1.

The number of animals required to conduct a bioassay for just one compound is significant. In addition, this design is not the most animal intensive that has been used, as some may employ 90 to 100 animals per gender and group and four or five dose levels.

The test substance is typically administered to animals daily (depending on the method of administration) over the course of their lifetimes. All the animals that die during the test interval or are killed at scheduled sacrifices are necropsied, gross pathology is performed, approximately 32 tissues are harvested, and micropathological evaluation on all tissues is completed by veterinary pathologists (9). All non-neoplastic and neoplastic findings are evaluated, statistically analyzed, and then interpreted by toxicologists and pathologists (7). The Carcinogenic Potency Project [Gold, LS. The Carcinogenic Potency Project. <http://potency.berkeley.edu/cpdb.html> (accessed November 11, 2002)] boasts that it contains the analysis of 5152 experiments on 1298 chemicals, including all of the NTP bioassays. Karstadt & Haseman (10) reported that, during the two decades

between 1976 and 1996, the NTP sponsored or conducted approximately 900 bioassays. In a similar time frame, the registrants of pesticides conducted an additional 620 bioassays [EPA. List of Chemicals Evaluated for Carcinogenic Potential. Communication Services Branch (7506C), Office of Pesticide Programs. <http://www.epa.gov/pesticides/carlist/> (accessed November 11, 2002)]. IARC [IARC Monographs Programme on the Evaluation of Carcinogenic Risks to Humans. <http://monographs.iarc.fr/> (accessed November, 2002)] lists a total of 879 agents, mixtures, and exposures that IARC has evaluated for carcinogenicity to humans.

Classification

As stated previously, all compounds identified as human carcinogens through epidemiologic studies have been demonstrated to be animal carcinogens. However, the converse of this argument, that all animal carcinogens are or can be expected to be human carcinogens, may not be true. The carcinogen classification schemes that have been developed by many agencies globally usually apply similar criteria, that is, the strength or weight of evidence (11–14). These guidelines call for the evaluation of mode of action data, along with human epidemiologic and animal cancer data, to assign the likelihood that a chemical may cause cancer in humans. The guidelines then call for a categorization that reflects whether or not a chemical is a known, probable, or possible human carcinogen; if the agent lacks carcinogenic activity; or if the agent cannot be classified.

Recently, the United States Congress has mandated the NTP to prepare a report that lists agents that are either known or reasonably anticipated to be human carcinogens. The NTP, like IARC, must use good scientific judgment and evaluate all relevant human data, animal data, mechanistic data, and conclusions regarding carcinogenicity in humans or experimental animals in listing a chemical (14).

The carcinogen classification schemes used by these various agencies are summarized in table 2.

Results

Over three decades of bioassay use allows us to step back and evaluate what it has given us. The results of evaluation and categorization of the findings of cancer bioassays by IARC and the EPA are summarized in table 3.

It can be seen that 43% of the chemicals reviewed by IARC were classified by IARC as known, probable, or possible human carcinogens, whereas 53% of the chemicals classified by the EPA fit the probable and

Table 1. Experimental design used by the National Toxicology Program for rodent cancer bioassays. ^a [From Chhabra et al (9)]

Parameter	Animals (N)	Species (N)	Gender (N)	Dose levels (N)	Totals
Exposure groups	60	2	2	3	720
Controls	60	2	2	1	240
Sentinel	15	2	2		60
Total					1020

^a Exposure duration: 104 weeks; interim evaluations: week 65 (10 animals/gender and group, killed for assessing chronic effects); toxicity end points: chronic toxicity, carcinogenicity.

Table 2. Carcinogen classification schemes. (IARC = International Agency for Research on Cancer, EPA = Environmental Protection Agency)

Category	Descriptor		
	IARC, 1992 (50) ^a	EPA, 1986 (11)	EPA, 1996 (12)
Carcinogenic to humans	Group 1	Category A	Carcinogenic to humans
Probably to carcinogenic to humans	Group 2a	Category B ₁ , category B ₂	Likely to be carcinogenic to humans
Possibly carcinogenic to humans	Group 2b	Category C	Suggestive evidence of carcinogenicity, but not sufficient to assess human carcinogenic potential
Unclassifiable as a human carcinogen	Group 3	Category D	Data are inadequate for an assessment of human carcinogenic potential
Probably not carcinogenic to humans	Group 4	Category E	Not likely to be carcinogenic to humans

^a The National Toxicology Program will use similar descriptors.

possible categories. Similarly, Haseman & Huff (8) found that about 51% of the 266 NTP studies they reviewed were positive in either the rat or mouse. Haseman (15) reported that, out of the 471 NTP studies, 250 (53%) produced carcinogenic effects in at least one gender-species group, 64 (14%) gave equivocal results, 150 (32%) were negative in all gender-species groups, and 7 (1%) were considered inadequate for evaluation. Thus about half of all chemicals tested have shown some carcinogenic activity that was judged to have some degree of significance to human health.

Rodent-human concordance

As previously described, both rats and mice have been used to assure the detection of carcinogenic potential. In essence, one species served as a "backup" for the other. Haseman & Huff (6) showed an overall concordance of 74% between rats and mice exposed to the same chemical. In addition, within a species, the response across genders had an 87% concordance for rats and 89% for mice. More amazingly, Haseman & Huff (6) concluded that using only male rats and female mice in these experiments would have achieved the same conclusion for 96% of the chemicals tested. Finally, the high interspecies correlation shown in these studies supports the view that extrapolation of carcinogenicity outcomes to other species, including humans, was validated. This conclusion supports the primary assumption that the carcinogenic response observed in animals can be directly translated to humans.

Gold et al (16) reviewed cancer bioassays for 738 chemicals (526 evaluated in rat bioassays and 412 in mouse bioassays) to examine concordance across species and target tissue sites. The top 10 target tissues for rodents are summarized in table 4.

Table 3. Classification of chemicals by the International Agency for Research on Cancer (IARC) and the Environmental Protection Agency (EPA).

Category	IARC ^a		EPA ^b	
	N ^c	Percent-age ^d	N ^c	Percent-age ^d
Carcinogenic to humans	174	10	0	0
Probably carcinogenic to humans	126	7	154	25
Possibly carcinogenic to humans	464	26	176	28
Unclassifiable as a human carcinogen	992	56	140	23
Probably not carcinogenic to humans	2	0.1	150	24
Total	1758		620	

^a IARC. IARC Monographs Programme on the Evaluation of Carcinogenic Risks to Humans. <http://monographs.iarc.fr/>. (accessed 11 November 2002).

^b United States Environmental Protection Agency. List of Chemicals Evaluated for Carcinogenic Potential. Communication Services Branch (7506C), Office of Pesticide Programs. <http://www.epa.gov/pesticides/carlist/> (accessed 11 November 2002).

^c Probable number of rodent bioassays used to classify.

^d Percentage of total bioassays.

Table 4. Ten most prevalent tumor sites in rodents [from Gold et al (16)].

Site	Positive chemicals	
	N	%
Rats		
Liver	206	39
Mammary gland	102	19
Kidney	86	16
Stomach	85	16
Hematopoietic system	56	11
Lung	55	10
Urinary bladder	49	9
Nasal cavity / turbinates	43	8
Ear / Zymbal's gland	41	8
Esophagus	36	7
Total chemicals	526	100
Mice		
Liver	231	56
Lung	116	28
Stomach	63	15
Vascular system	59	14
Hematopoietic	53	13
Kidney	26	6
Mammary gland	24	6
Thyroid gland	18	4
Urinary bladder	12	3
Uterus	12	3
Total chemicals	412	100

Of the 10 most frequent tumors detected in rats or mice, seven are common to both (70% concordance), indicating a relatively high degree of correlation between rodent species.

Liver tumors are the most common response in rodent cancer bioassays (16, 17). In fact, a deliberate decision was made 30 years ago to choose strains of rodents exhibiting high liver sensitivity to known carcinogens

(17). However, as presented in table 5, the age-adjusted incidence of liver tumors in humans is 2.4/100 000 (18). Mammary gland tumors in rats and mice were approximately 19 000/100 000 and 6000/100 000, respectively, as opposed to breast cancer in humans at 134.1/100 000. Although there was 70% concordance for the most prevalent sites between rat and mouse (table 4), there was only 30% site concordance between rats and humans and 40% between mice and humans (table 5).

Since the bioassay is used to predict the likelihood of cancer occurring in humans, a useful comparison is the incidence rates for cancer in humans. Table 5 shows the incidence and mortality for the top 10 human cancers. Whereas these are not cancers presumed to be induced or elevated by chemicals, as in the rodent cancer table (table 4), the incidence rates give us some idea of how rodent cancer compares to human cancer rates.

Table 5. Ten most prevalent tumor sites in humans (19).

Site	Incidence per 100 000 ^a	Deaths per 100 000
Prostate (male)	165.4	12.6
Breast (female)	134.1	28.8
Lung	62.9	57.7
Colon and rectum	53.7	21.7
Corpus uteri (female)	24.2	2.1
Urinary bladder	20.0	4.4
Lymphoma ^b	19.1	8.7
Ovary (female)	16.8	9.0
Skin melanoma	16.3	2.7
Oral cavity and pharynx	11.1	3.0

^a Age-adjusted to 2000 US population.

^b Non-Hodgkin's lymphoma.

Table 6. Summary of results of the National Toxicology Program for 266 bioassays of male and female rats and mice [Haseman & Huff (6)]. (+ = positive in that species and gender, – = negative in that species and gender)

Model	Possible outcomes				Number of studies
	Fisher 344 rats		B6C3F1 mice		
	Male	Female	Male	Female	
4	+	+	+	+	38
3	+	+	+	—	1
	+	+	—	+	10
	+	—	+	+	5
	—	+	+	+	5
2	+	+	—	—	16
	+	—	+	—	1
	+	—	—	—	4
	— ^a	+	+	— ^a	1
	—	+	—	+	2
	—	—	+	+	23
1	+	—	—	—	13
	— ^a	+	—	— ^a	3
	— ^a	—	+	— ^a	7
	—	—	—	+	6
0	—	—	—	—	131

^a Chemical that would not have been found carcinogenic had only the male rat and female mouse models been used.

Only 2 of the top 10 human cancers are concordant with appearance in rodents.

The fact that there is minimal concordance between the target organs most prevalent in the rodent to those in humans is at least worrisome. This concern may have contributed to Karstadt & Haseman's (10) decision to review bioassay databases for mouse liver, rat thyroid follicular cell, bladder and kidney (male rat), forestomach (mouse and rat, gavage route), and lung (mouse and rat, inhalation of particles). These are organ sites in the top 10 sites for rodents (see table 4) for which the mode of action is thought to be partially known and is considered to be irrelevant to humans. These investigators wished to determine the effect of discounting the tumors on the classification of chemicals as carcinogens in animals; thus they examined both the carcinogenicity data for chemicals studied in NTP bioassays or the ones reviewed in IARC monographs. Karstadt & Haseman (10) found that deleting rodent liver tumors had a significant impact on evaluations of carcinogenicity in animals, that is, 56 of 234 (24%) chemicals determined to be carcinogenic in the NTP bioassay program would no longer be considered carcinogenic. In addition, 102 of 234 (44%) would have weaker evidence of carcinogenic effects. For the IARC database, 33 of 361 (9%) chemicals determined by IARC to have "limited" or "sufficient" evidence of carcinogenicity would no longer be considered carcinogenic, and 119 of 361 (33%) would have weaker evidence of carcinogenic effects. Thus liver tumors in rodents drive a substantial portion of the decision to call a chemical "carcinogenic", yet liver tumors are not prevalent in the United States population.

After two decades of testing, these results lead to either of two diverging conclusions, either the early premise holds that site concordance is not necessary to simply identify compounds with carcinogenic potential or the bioassay has failed to properly alert us to cancers of human relevance.

Strain comparison

Along with the question of organ concordance between species, there comes the less-examined relationship of strain differences in characterizing the carcinogenic response. In an attempt to reduce the number of animals used and evaluated in the bioassay, scientists at the NTP examined the possibility of eliminating some of the four models that they used in their bioassay (ie, the male rat, female rat, male mouse and female mouse). Haseman & Huff (6) concluded that, if they had only used male F-344 rats and female B6C3F1 mice, they would be able to detect over 90% of the tumors observed from using both male and female rats and mice of these strains. This claim is illustrated in table 6, where 266 chemicals were examined in two species and two genders (four models).

Table 6 shows that, of the 266 chemicals examined, there were 135 chemicals (~51%) tested that had a positive response in at least one or more of the four NTP models. There were 131 chemicals that did not yield a positive response in any of the NTP models (~49%). Of the 135 chemicals that were determined to be carcinogenic in the bioassay, 124 chemicals (92%) would have been found to be positive in either or both the male Fischer 344 rat and the female B6C3F1 mouse. Eleven chemicals (8%), which were positive in models in either or both the female Fisher 344 rat and male B6C3F1 mouse, would have been missed if only the male rat and the female mouse models were considered.

Interestingly, a different conclusion can be gained if we perform a similar exercise using the results from bioassays conducted for 46 chemicals of a certain former crop protection chemical company. Industry traditionally uses other strains of rodents for their bioassays than government laboratories do. Instead of the F-344 rat, the Sprague-Dawley rat is used; instead of B6C3F1 mice, CD-1 mice are used. If the strains of rodents are switched, much different results are obtained, as presented in table 7.

Of the 46 chemicals evaluated, 24 of the bioassays were positive in one or more of the four models (~52%). However, if only the male rat and the female mouse models are used, as proposed in the NTP scenario, 14 chemicals (64%), which were found to be animal carcinogens in the rodent bioassay models, would have been missed. Therefore, the strain of rodent used in evaluating the oncogenic potential of a chemical in a bioassay can markedly affect the ability to effectively assess the carcinogenic potential of a chemical in animals.

Notably, the tumor types missed by the two NTP models (male Fisher-344 rat and female B6C3F1 mouse) were mammary gland tumors for eight chemicals observed in the Sprague-Dawley female rat, and six chemicals produced liver tumors in the male CD-1 mice. As the mode of action is known for the female rat model and is known to be irrelevant to man (20) and the mode of action for liver tumors in male mice has been at least mildly challenged regarding human relevance (21), it can be questioned whether anything important was actually missed. Therefore, the gender and strain of rodent used in the cancer bioassay can have a significant impact on the cancer outcome observed.

We have seen that the evidence for organ sites does not necessarily correlate well from rodents to humans, and the selection of the rodent strain chosen may alter the conclusion, but there are other factors that need to be considered because they may have an impact on the primary assumptions. Half of the six primary assumptions are associated with dose-response, namely, the relationship between dose and cancer response is linear, the slope of the dose-response relationship at low doses

can be derived from data at high doses, and large doses, including the MTD, must be employed to insure reasonable power for tumor detection (7). Thus the effect of dose on cancer detection deserves careful consideration.

Effect of dose level selection

Three of the six basic assumptions concerning the interpretation of bioassay results made by Haseman (7) involve the effect of dose on response to carcinogens (ie, the relationship between dose and cancer response is linear, the slope of the dose-response relationship at low doses can be derived from data at high doses, and large doses, including the MTD, must be employed to insure reasonable power for tumor detection). All three of these assumptions assume that an investigator can confidently use a linear low-dose extrapolation procedure to accurately estimate low-dose human cancer risk directly from the rodent bioassays regardless of whether the doses selected are low, high, or even an overdose. The MTD concept, as it is empirically set for a given chemical in a bioassay, often embodies overdosing.

The impact of administering an excessive MTD of chemicals on the tumorigenic response in the rodent bioassay cannot be accurately estimated without careful reviews of the completed bioassays; however, it is known to be significant (22, 23).

Haseman & Seilkop (24) examined how the association between rodent carcinogenicity and MTD was evaluated in 326 long-term carcinogenicity studies with mice and rats; they focused on experimental outcomes.

Table 7. Summary of the results of bioassays of company X for the Sprague-Dawley rat and CD-1 mouse for 46 crop protection chemicals. (+ = positive in that species and gender, - = negative in that species and gender)

Model	Possible outcomes				Number of studies
	Sprague-Dawley rats		CD-1 mice		
	Male	Female	Male	Female	
4	+	+	+	+	0
3	+	+	+	−	0
	+	+	−	+	0
	+	−	+	+	0
	−	+	+	+	0
	2	+	+	−	−
2	+	−	+	−	0
	+	−	−	−	0
	−	+	+	−	0
	−	+	−	+	8
	−	−	+	+	0
	1	+	−	−	−
1	− ^a	+	−	− ^a	8
	− ^a	−	+	− ^a	6
	−	−	−	+	0
	0	−	−	−	−

^a Chemical that would not have been found carcinogenic had only the male F-344 rat and female B6C3F1 mouse models been used.

These investigators found that the MTD was not a reliable predictor of whether or not a chemical would be a rodent carcinogen. Their investigation confirmed that comparisons of carcinogenic potencies based only on positive studies may result in artificially elevated estimates of the underlying association between chemical toxicity and rodent carcinogenicity. Gaylor & Gold (25) concluded that this finding was, in part, due to the relatively narrow range of doses employed in 2-year rodent bioassays and the limited range of statistically significant tumor incidence rates.

Chronic, accelerated cell division is plausible as the major reason for more than half of the studied chemicals being classified as carcinogens when tested at the MTD in standard rodent cancer bioassays (26–28). As currently conducted, rodent bioassays provide inadequate data to estimate human risk at low doses. Sixty percent of the chemicals tested in both rats and mice are carcinogenic; even among the known nonmutagens, 49% are positive (among the mutagens, 78% are positive) (28). The high positive rate is consistent for synthetic chemicals, natural chemicals (99.9% of the chemicals humans are exposed to are natural) (28–30), natural pesticides, and chemicals in roasted coffee (27, 31). The carcinogenicity of constituents in roasted coffee, as identified by the rodent bioassays, is presented in table 8.

Therefore, as excessive doses have been routinely used in rodent bioassays, the human relevance of a tumor observed at these high levels of exposure has to be challenged. Even if one assumes a linear relationship between dose and cancer response, the slope of the dose–response relationship at low doses can be derived from data derived at the excessively high levels.

Refinements proposed for the bioassay

We have considered the fact that organ-site evidence does not translate well from rodents to humans, the selection of strain of rodent may alter the outcome, the relationship between dose and cancer response may not be linear, and excessive dosing may enhance the power of the bioassay but lead to false positives. Hence the six

Table 8. Carcinogenic hazard of natural constituents of roasted coffee as assessed by the rodent bioassay [Ames & Gold (27)].

Assessment ^a	Coffee constituent
Positive (N=19)	Acetaldehyde, benzaldehyde, benzene, benzofuran, Benzo (a) pyrene, caffeic acid, catechol, 1,2,5,6-dibenazanthracene, ethanol, ethylbenzene, formaldehyde, furan, furfural, hydrogen peroxide, hydroquinine, limonene, styrene, toluene, xylene
Negative (N=8)	Acrolein, biphenyl, choline, eugenol, nicotinamide, nicotinic acid, phenol, piperidine
Uncertain	Caffeine

^a To be tested: about 1000 chemicals.

major assumptions that were made when the format of the cancer bioassay was first established (7) would appear to not have weathered the test of time as leading to relevant findings for humans that are meaningful in human risk assessment. Therefore, a new paradigm should be considered that incorporates new molecular biological methods, establishes a plausible mode of action that can be understood, applies across species, genders, and ages while using pharmacokinetic and dynamic principles that permit the comparison of the rodent model with the human one and reduce the number of animals utilized in these tests.

Incorporating new molecular biological methods

After more than five decades of time-consuming, resource-intensive, animal-intensive, and less than all-predictive rodent cancer bioassays being conducted, it is becoming increasingly apparent that there has to be a better way of assessing human carcinogenic risk and incorporating more information on mechanisms and modes of action into the risk assessment process (32). It has been estimated that the world's total resources for the activity of rodent bioassays could only perform a few hundred bioassays in a 3-year period. As was pointed out by Fung et al (33) in 1995, over 75 000 chemicals in commercial use have yet to be evaluated. The question that now confronts scientists is how we can create alternative approaches that accurately and efficiently capture the cancer process and reduce the unnecessary use of animals (34). The following is a discussion of approaches that may result in the reduction of animal use.

Ames assay. Over 40 years ago, there was great excitement concerning one possible alternative approach, mutagenicity testing (35). In the time that followed, significant effort was spent on the standardization and validation of these tests, as well as on accepting a battery of these short-term in vitro and in vivo tests (36, 37). Clearly, this approach had significant promise in reducing the number of animals used, in reducing time and costs, and in enhancing predictability, as it was found that most human carcinogens are genotoxic (mutagenic), but not all genotoxic agents have been shown to be carcinogenic in humans (38). Despite a well-defined correlation between mutagenic, or DNA-reactive potential, and carcinogenic potential, only about 70% of the mutagens have been determined to be carcinogenic in the NTP database (39). Although the correlations have become weaker over the past 20 years, and with them the predictive value of short-term tests, such tests still provide vital information for identifying and understanding the mechanisms involved in carcinogenicity (40). Despite the possible weaknesses in the genotoxic assessment approach, it may be equal or better than the

rodent cancer bioassay in detecting chemicals that may cause cancer in humans.

Structure-activity relationships. An additional alternative method that demonstrated some promise was the evaluation of chemicals by structure-activity relationships (SAR) as another alternative to the 2-year bioassay (41, 42). SAR models are generally built to include many of the parameters (eg, log P and organic media) necessary for modeling in vivo activity, both linear and nonlinear regression models, and generic algorithms to select an optimal set of molecular descriptors for modeling and predicting activity [Welch WJ. Computational tools for the prediction and classification of estrogenic compounds. Available from URL: <http://cfpub.epa.gov/ncer/abstracts/index.cfm/fuseaction/display.abstractDetail/abstract/169/report/0> (accessed 17 November 2002)]. The algorithms can be built upon chemical or physical properties of the molecule, or actually built from a database of bioassays containing the biological response for classes of chemicals. If the actual experimental results of a database of NTP bioassays are used, it should be well understood that the reliability of the estimate is no better than the accuracy of the reliability of the database used to construct the algorithms for the model.

In 1996, the NTP ran a test demonstration to evaluate the predictability of several commercial and knowledge-based models in predicting the outcomes of rodent carcinogenicity tests (43). In the first set of 44 chemicals of unknown carcinogenic potential for which NTP bioassays were run, 13 participant modelers were given the chemical structures and asked to predict with their specific model for carcinogenicity (positive or negative response) for each of the chemicals, without the modelers' a priori knowledge of the outcome of the bioassays. Since 9 of the 44 chemicals were equivocal, only the remaining 35 chemicals were used to estimate the overall predictive success of the 13 modelers. The results of this assessment are presented in figure 1.

From this demonstration project, modelers 1–3 and 7 showed some promise, but none of the 13 modelers provided predictive estimates that were much better than a toss of a coin. SAR has not yet achieved the level of reliability needed for it to be used exclusively for establishing the carcinogenic potential of a chemical.

Genetically modified animals. Additional alternatives lead back to in vivo tests. Advances in genetic engineering have created opportunities for improved understanding of the molecular basis of carcinogenesis. Investigators can produce mice of unique genotypes and phenotypes that afford insights into the events and mechanisms responsible for tumor formation (44).

Eastin et al (45) reported that researchers of the National Institute of Environmental Health Sciences are

exploring the utility of genetically altered mice to study mechanisms of carcinogenesis. Two of these mouse models, the Tg.AC (carrier of an activated mouse *H-ras* oncogene) and the p53+/- (heterozygous for the wild-type tumor suppressor gene Trp53), have genetic alterations that appear to hasten their expression of chemically induced tumors. In fact, the first set of 11 chemicals tested in genetically altered mice compared with previous findings in the traditional 2-year rodent assays and the literature on human tumor findings (45).

This alternative approach would eliminate the need to perform a full lifetime mouse carcinogenicity study, but replace full mouse bioassay using approximately 1000 animals for both genders with a study of 100 to 200 transgenic mice of one gender; the treatment period would be reduced from 1.5 years to 26 weeks (46). Clearly, these models would appear to have the potential to serve as more rapid and less expensive test systems with which to identify carcinogens than the standard rodent bioassay.

In fact, the Food and Drug Administration has agreed that available information indicates that there is sufficient experience with some in vivo transgenic rodent carcinogenicity models to support their application as complementary second species studies, in conjunction with a single 2-year rodent carcinogenicity study (47). The optional substitution of a second 2-year rodent carcinogenicity study with an alternative study, such as an in vivo transgenic carcinogenicity study, is part of the International Conference on Harmonization guidance S1B: Testing for Carcinogenicity of Pharmaceuticals (48).

Shortening the bioassay. Another practical consideration may be to shorten the duration of the bioassays. Davies et al (49) studied the IARC monographs (volumes 1–70) to determine the time of the onset of treatment-related tumors in long-term rodent studies for chemicals classified by IARC as having sufficient evidence of carcinogenicity in animals. Their analysis excluded studies on metals and their salts, studies on particulates, studies by parenteral routes of administration that resulted in tumors only at the site of exposure, and studies that did not equate to the current standard long-term

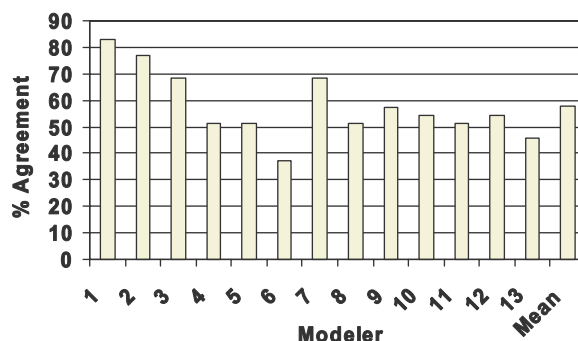


Figure 1. Summary of 13 sets of modeler predictions made for 35 chemicals evaluated using cancer bioassays of the National Toxicology Program.

rodent carcinogenicity bioassay. Two hundred and ten chemicals were reviewed. Davies et al (49) found that treatment-related tumors were first apparent within 12 months for 66% of the chemicals and that studies of longer than 18 months were necessary for only 7%. The list of chemicals included IARC group 1 chemicals (human carcinogens) that were detected in animals within 18 months; most were found within 12 months. Most of the tumor types that required more than 18 months for detection were of dubious relevance to risk assessment for human health. Therefore, the study duration could be cut in half and not miss any important chemically related tumor formation.

Establishing a plausible mode of action that can be understood and applied across species, genders and ages

In hazard identification and subsequent hazard characterization, the primary question is whether the effects observed in animal studies can be expected in humans. Because of the grave consequences of cancer, certainty about cancer causation has always been a paramount concern. Assessing the relevance of animal tumors to human health reliably demands an understanding of the pathobiology that underlies the carcinogenic effect (13, 50–52). Recent scientific research has yielded a substantial body of data on carcinogenic modes of action in laboratory animals. Studies in laboratory animals and, to a less extent in humans, firmly establish that cancer arises through genetic and nongenetic events and that informed use of information about a carcinogen's mode of action could increase confidence in conclusions about the hazard and risk potential of environmental chemicals for humans.

To address this issue, the US EPA and the International Programme on Chemical Safety (IPCS) proposed guidance for assessing the weight of evidence for modes of action for chemical carcinogens in animals as an aid in determining whether animal tumors are relevant for human cancer risk assessment [EPA. Guidelines for carcinogen risk assessment: risk assessment forum: SAB review draft. Washington (DC): US Environmental Protection Agency; 1999. (Available at URL: <http://www.epa.gov/ncea/raf/crasab.htm>); 53]. The proposed guidance for mode of action focuses on identifying "key events", generally described as "measurable effects that are critical to the induction of tumors, as hypothesized in the postulated mode of action", for use in evaluating the mode of induction of these tumors. Thus the actual relationship of findings in animals to human health depends on knowing the key events by which an agent causes cancer in animal models and whether or not those key events would plausibly occur in humans.

A workgroup from the International Life Sciences Institute (ILSI) and the Risk Sciences Institute (RSI) continued the development of the EPA/IPCS guidance

by examining the data on mode of action for six chemically induced tumor types. From these case studies, the committee recommended a framework (54) for more fully using information on mode of action to analyze the human relevance of animal tumors in the hazard characterization phase of the risk assessment process.

The framework is comprised of three questions and a concordance table. The three questions are a stepwise analysis of data that establishes (or does not) the plausibility of animal effects occurring in humans. The concordance table allows a side-by-side comparison of key events in animals and humans and clearly shows whether or not there is a plausible relationship (concordance). The questions are as follows: (i) is the weight of evidence sufficient to establish the mode of action in animals (postulated mode of action, identification of key events, and application of EPA/IPCS guidance on mode of action (based on Bradford-Hill criteria), (ii) is the animal mode of action plausible for humans (concordance table and analysis and statement of confidence), and, taking into account kinetic and dynamic factors, is the animal mode of action plausible for humans?

Two critical features characterize this stepwise analysis. One is the careful and systematic definition of key events in the mode of action and the other is the concordance table.

Key events are identified in question 1, for which data is evaluated on a weight-of-evidence basis in the style of the Bradford-Hill criteria. The Bradford-Hill criteria rely on a systematic consideration of data that either strengthen or weaken a postulated mode of action. In the first instance, the question is whether the agent under investigation is responsible for the tumorigenic effects observed in animals, with comparable questions about related information on mode of action. In this regard, the Bradford-Hill criteria (citation) for causation serve as a general guide. Epidemiologists have used these criteria to rigorously evaluate data, and the IPCS has recommended their use as a means to identify causality. Causation generally depends on the correspondence of temporal relationships, specificity, biological plausibility, and overall coherence of the database. Other considerations include consistency, the strength of the association, and dose-response relationships. Not all criteria must be fulfilled to find causation, and the findings vary in different situations. Nonetheless, this rigorous process is fundamental to the decision as to whether or not a mode of action has been identified.

The concordance table forces the assessor to carefully list the key events in animals and attempt to align them with similar processes in humans. If carcinogenic key events occur in both animals and humans, then that agent would likely cause cancer in humans. If, on the other hand, concordance cannot be established for key events in animals and humans, then the mode of action

is not likely to be operative in humans, and a carcinogenic outcome is not plausible.

Adherence to this disciplined analytical approach to the mode of action of a carcinogenic compound provides a reliable assessment of the relevance to human health with respect to the tumorigenic effect on animals. It also provides a rational guide to the development of data to understand carcinogenic modes of action in animals and humans.

Developing pharmacokinetic and dynamic data that permit the comparison of the rodent model to the human mode

Models generally consider risk to be a function of the hazard (toxicity) and exposure (dose). That function is best described by the dose response of the toxic effect (22). For any risk assessment system to be effective, it should consider the dose-response relationship as depicted in figure 2.

The administration of very high doses to animals (eg, the MTD) to maximize hazard identification can produce nonlinear toxicokinetics due to the saturation of the metabolism or excretion of the compound. Saturation phenomena often produce nonlinear dose curves, and any risk assessment system should be able to address such effects. Further complications include the increased involvement of normally minor metabolic pathways.

A default assumption in most low-dose extrapolations is that there is a linear relationship between dosage and target organ exposure to the active chemical entity. Sumner & Stevens (22) presented several examples for the impact of the MTD on the kinetics of saturable processes and how risk assessment can be impacted. One such case of saturable processes is represented by altered metabolism at high levels of administration or at the dose equal to the MTD. A case history for sulfonylurea is used for this example. Sulfonylurea is moderately soluble in water and has a molecular weight of 468.3, making it a candidate for biliary secretion (55). This product is not particularly acutely toxic; it is not a mutagen, a teratogen, or a reproductive toxin. It is not tumorigenic to rats. It has not been demonstrated to be a potential oncogen, but, at doses above the MTD, liver tumors were found in mice (56). At doses below the MTD, the compound is metabolized by the hydrolysis of sulfonylurea and oxidation, as well as by conjugation. Feeding levels above 3000 ppm resulted in a new metabolite that was found in mice feces. This compound seems to demonstrate activation of a secondary pathway of degradation at these high doses.

A histogram of the chromatography of the radioactive fecal metabolites of sulfonylurea is presented in figure 3. This histogram shows that early eluting metabolites may change quantitatively with increasing dose

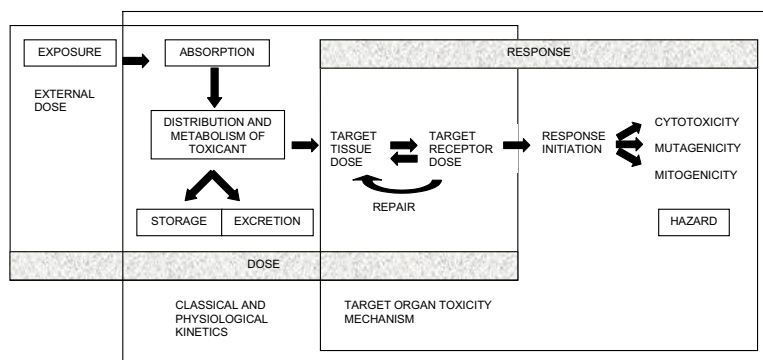


Figure 2. Depiction of pharmacokinetics and pharmacodynamics incorporated into the toxicologic process.

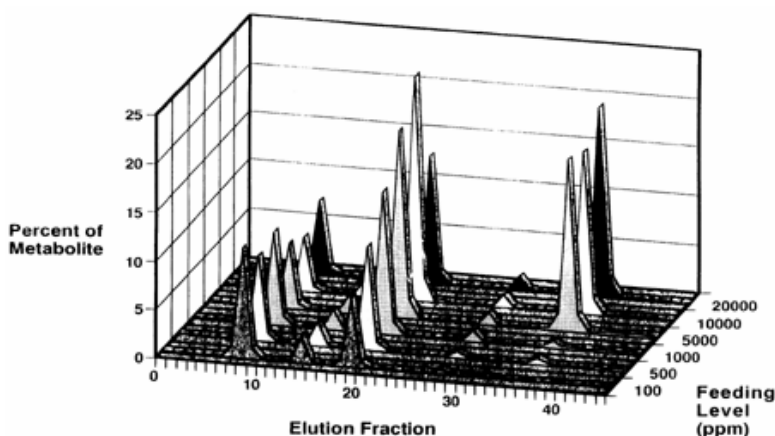


Figure 3. Radiochromatogram of sulfonylurea metabolites in feces of male mice.

but that the late eluting metabolite appears only at high doses. Interestingly, this late peak, eluting at fraction 36, seems to be less polar than the parent that elutes at fraction 28 in this reverse-phase column. It is also interesting to note that the metabolite seems to form only after repeat dosing, not with a single dose. The metabolite was identified after mice were fed radiolabeled sulfonylurea for 28 days. The high-dose metabolite was found to be a dimer of the original compound with a ring difluoromethoxy group replaced with a disulfide bridge between two molecules of the test material.

In addition to clarifying the metabolite milieu at high-dose levels versus low, pharmacokinetic data may also help resolve a fundamental discrepancy between the long-term bioassay and human health risk assessments. Animal studies are performed at a constant dose level for a fixed duration (2 years). However, human exposures are intermittent and may fluctuate from high to low. Kinetic data would provide a rational basis for bridging between constantly dosed animal studies and intermittently exposed humans.

Choosing dose levels so that the highest doses tested do not greatly exceed potential exposure scenarios for humans

The classic assumption in the early days of long-term bioassays was that if a compound can be pushed into an animal for its lifetime at the maximally tolerated dose (ie, without killing the animal) then the investigator would have assessed the full range of potential toxicologic outcomes. Whereas we used high doses to increase the sensitivity of hazard identification, unfortunately, many of the problems in toxicologic studies would arise from trying to use high-dose effects on animals to predict potential effects for low-dose exposures in humans.

As our understanding of carcinogenic processes improves, we are realizing that former assumptions about linear low-dose extrapolation, particularly for nongenotoxic compounds, are flawed. As with any toxicologic effect, threshold-based mechanisms underlie the carcinogenic process and what happens at high-dose levels would not necessarily occur at low-dose levels.

Table 9. Advantages of the genetically modified mouse as a cancer bioassay in comparison with the standard bioassay of the B1C3F1 mouse.

Mouse strain	Duration of exposure (months)	Animals required (N)	Age-related diseases	Spontaneous tumors (at end of study)	End of study survival (controls) (%)
B6C3F1	18–24	200	Yes	Yes	59–79
P53 +/-	6	60	No	No	>90
Tg.AC	6	60	No	No	<90

Nonetheless, we still conduct studies that require exposures at the MTD. Perhaps we should either reconsider the relevance of these findings at high-dose levels or modify dose-level requirements.

Reducing the number of animals used in tests

As shown in table 1, over 1000 animals are used in each long-term bioassay. The reasons for using this many animals include statistical power viability and reliable sensitivity to detect tumors. Nonetheless, our quest to reveal the cancer-causing potential of chemicals should take into consideration the opportunity to limit the number of animal lives required to achieve that quest. In 1959, the efforts of Russell & Burch (57) identified the “3Rs” for animal welfare, replacement, reduction, and refinement.

Relevance is a key feature of alternative methods to using animals and is the extent to which a test correctly measures or predicts the biological effect of interest (58). In the case of cancer bioassay, the biological end point of interest is cancer in humans. The continued use of rodent cancer bioassays as a fundamental pillar of carcinogenicity testing does not lead any closer to the realization of this objective. As previously discussed, these studies are performed at a high cost in terms of time, financial resources, and animals. Compounding these expenses is the fact that many of the rodent cancer bioassays yield equivocal or irrelevant results in regard to humans (59). Clearly there is a critical need to come up with an alternative model for the rodent cancer bioassay that is relevant, reliable, rapid, and reasonable in cost. If we accept the basic concerns of ICVAM (58) and ECVAM (34) for alternative tests for the rodent cancer bioassay, we will look to tests that embody the basic concept of the replacement, reduction, or refinement of the use of animals. We have considered wonderful replacement options of genotoxicity testing (Ames test) and structure activity relationships in the replacement area. Although neither approach requires animals, unfortunately, neither can be touted as replacements for the cancer bioassay as they fail with respect to either or both relevant and reliable criteria (60). Therefore, we must seek additional supplemental approaches to cover the reduction or refinement criteria. In these areas, we have considered the genetically modified mouse. The genetically modified P53 +/- and Tg.AC mouse models do fit in both reduction and refinement areas as the number of animals are reduced as presented in table 9.

Both the P53 +/- and Tg.AC reduce the duration of exposure by 3- to 4-fold and the number of rodents required by 31- to 33-fold. These assays meet the relevant, rapid, and reasonable cost criteria; the verdict still may be out as to reliability. The genetically modified

mouse bioassay may add great value as a part of the tierwise testing strategy in cancer assessment (61).

Developing an internationally harmonized, tierwise testing strategy

Overall, the long-term bioassay has served us well in verifying or identifying carcinogenic compounds. However, as discussed in this paper, there are many areas that are ripe for improvement. To address this concern, the International Life Sciences Institute (ILSI) has convened a panel of international experts in toxicity testing to define, from first principles, how best to establish the toxicologic characteristics of compounds. The mission statement of this group is, in part, to provide a mechanism for reaching consensus across sectors (industry, government, academia, and nonprofit) on scientific issues of concern relating to the current safety assessment process for crop protection chemicals. One of the chief aims of the workgroup was to define a process for testing chemicals in a tierwise fashion, such that one test result leads to a second, third, and so forth, leading up to a long-term study to address carcinogenic potential from long-term exposure. The field of toxicology and, indeed, epidemiology awaits the output of this group in the ongoing quest to properly identify the carcinogenic potential of chemicals.

References

- Potter M. Percivall Pott's contribution to cancer research. In: National Cancer Institute (NCI). Conference: Biology of Cutaneous Cancer. Bethesda (MA): NCI; 1962. Monograph no 10:1-13.
- Yamagiwa K, Ichikawa K. Experimental study of the pathogenesis of carcinoma. *J Cancer Res* 1915;3:1.
- Wilbourn JD, Haroun L, Vainio H, Montesano R. Identification of chemicals carcinogenic to man. *Toxicol Pathol* 1984;12(4):397-9.
- Tomatis L, Aitio A, Wilbourn J, Shuker L. Human carcinogens so far identified. *Jpn J Cancer Res* 1989;80(9):795-807.
- Huff J, Cirvello J, Haseman J, Bucher J. Chemicals associated with site-specific neoplasia in 1394 long-term carcinogenesis experiments in laboratory rodents. *Environ Health Perspect* 1991;93:247-70.
- Haseman JK, Huff JE. Species correlation in long-term carcinogenicity studies. *Cancer Lett* 1987;37:125-32.
- Haseman JK. Statistical issues in the design, analysis and interpretation of animal carcinogenicity studies. *Environ Health Perspect* 1984;58:385-92.
- Haseman JK, Huff JE, Zeiger E, McConnell, EE. Comparative results of 327 chemicals carcinogenicity studies. *Environ Health Perspect* 1987;74:229-35.
- Chhabra RS, Huff JE, Schwetz BS, Selkirk J. An overview of prechronic toxicity / carcinogenicity experimental study designs and criteria used by the National Toxicology Program. *Environ Health Perspect* 1990;86:313-21.
- Karstadt M, Haseman JK. Effect of discounting certain tumor types/sites on evaluations of carcinogenicity in laboratory animals. *Am J Ind Med* 1997;31(5):485-94.
- US Environmental Protection Agency. Guidelines for carcinogen risk assessment. *Fed Regist* 1986;51(185):33992-34003.
- US Environmental Protection Agency (USEPA). Proposed guidelines for carcinogen risk assessment. Washington (DC): USEPA; 1996. EPA/600/P-92/003C.
- Vainio H, Magee PN, McGregor DB, McMichael AJ, editors. Mechanisms of carcinogenesis in risk identification. Lyon: International Agency for Research on Cancer; 1992.
- National Toxicology Program (NTP). Report on carcinogens. 8th edition. Research Triangle Park (NC): NTP; 1998.
- Haseman JK. Using the NTP Database to assess the value of rodent carcinogenicity studies for determining human cancer risk. *Drug Metabol Rev* 2000;32(2):169-86.
- Gold LS, Manley NB, Slone TH, Ward JM. Compendium of chemical carcinogens by target organ: results of chronic bioassays in rats, mice, hamsters, dogs, and monkeys [review]. *Toxicol Pathol* 2001;29(6):639-52.
- Gold LS, Slone TH, Manley NB, Bernstein L. Target organs in chronic bioassays of 533 chemical carcinogens. *Environ Health Perspect* 1991;93:233-46.
- El-Serag HB, Mason AC. Rising incidence of hepatocellular carcinoma in the United States. *N Eng J Med* 1999;340(10):745-50.
- National Cancer Institute (NCI) [homepage on the Internet]. SEER cancer statistics review: 1973-1999. Pontiac (MI): NCI; 2002. Access: [http://seer.cancer.gov/csr/1973_1999].
- Stevens JT, Breckenridge CB, Wetzel LT, Thakur AK, Liu C, Werner C, et al. A risk characterization for atrazine: oncogenicity profile. *J Toxicol Environ Health A* 1999;56:69-109.
- Williams GM. Chemicals with carcinogenic activity in the rodent liver; mechanistic evaluation of human risk. *Cancer Lett* 1997;117(2):175-88.
- Sumner DD, Stevens JT. Pharmacokinetic factors influencing risk assessment: saturation of biochemical processes and cofactor depletion. *Environ Health Perspect* 1994;102 suppl 11.
- Van Gemert M, Stevens JT, Cardona RA, Hurt S. Implications of international differences in dose selection: chapter 3. In: Foran J, editor. ILSI monograph on dose selection. Washington (DC): Risk Science Institute; 1997. p 95-114.
- Haseman JK, Seilkop SK. An examination of the association between maximum-tolerated dose and carcinogenicity in 326 long-term studies in rats and mice. *Fundam Appl Toxicol* 1992 Aug;19(2):207-13.
- Gaylor DW, Gold LS. Regulatory cancer risk assessment based on a quick estimate of a benchmark dose derived from the Maximum Tolerated Dose (MTD). *Regul Toxicol Pharmacol* 1998;28:222-5.
- Ames BN, Shigenaga MK, Gold LS. DNA lesions, inducible DNA repair, and cell division: three key factors in mutagenesis and carcinogenesis. *Environ Health Perspect* 1993;101 suppl 5: 35-44.
- Ames BN, Gold LS. Chemical carcinogenesis: too many rodent carcinogens. *Proc Natl Acad Science U S A* 1990;87:7772-6.
- Gold LS, Manley NB, Slone TH, Garfinkel GB, Rohrbach L, Ames BN. The fifth plot of the Carcinogenic Potency Database: results of animal bioassays published in the general literature through 1988 and by the National Toxicology Program through 1989. *Environ Health Perspect* 1993;100:65-135.
- Gold LS, Slone TH, Stern BR, Manley NB, Ames BN. Rodent

- carcinogens: setting priorities. *Science* 1992;258:261–5.
30. Ames BN, Profet M, Gold LS. Dietary pesticides (99.99% all natural). *Proc Natl Acad Science U S A* 1990;87: 7777–81.
31. Gold LS, Bernstein L, Magaw R, Slone TH. Interspecies extrapolation in carcinogenesis: prediction between rats and mice. *Environ Health Perspect* 1989;81:211–9.
32. Sills RC, French JE, Cunningham ML. New models for assessing carcinogenesis: an ongoing process. *Toxicol Lett* 2001;120(1–3):187–98.
33. Fung VA, Barrett JC, Huff J. The carcinogenesis bioassay in perspective: application in identifying human cancer hazards. *Environ Health Perspect* 1995;103:680–3.
34. Combes R, Balls M, Curren R, Fischbach M, Fusenig N, Kirkland D, et al. Cell transformation assays as predictors of human carcinogenicity: the report and recommendations of ECVAM Workshop 39. Available from URL: <http://altweb.jhsph.edu/publications/ECVAM/ecvam39.htm>. Reprinted with minor alterations from *Altern Lab Anim* 1999;27(5): 745–67.
35. Ames BN, Whitfield HJ Jr. Frameshift mutagenesis in *Salmonella*. *Cold Spring Harb Symp Quant Biol* 1966;31:221–5.
36. Ashby J, Waters MD, Preston J, Adler ID, Douglas GR, Fielder R, et al. IPCS harmonization of methods for the prediction and quantification of human carcinogenic/mutagenic hazard, and for indicating the probable mechanism of action of carcinogens. *Mutat Res* 1996;10;352(1–2):153–7.
37. Water MD. Development and impact of the Gene-Tox-Program, genetic activity profiles, and their computerized databases [review]. *Environ Mol Mutagen* 1994;23 suppl 24:67–72.
38. Albertini RJ, Anderson D, Douglas GR, Hagmar L, Hemminki K, Merlo F, et al. IPCS guidelines for the monitoring of genotoxic effects of carcinogens in humans: International Programme on Chemical Safety. *Mut Res* 2000;463(2):111–72.
39. Tennant R. What is a tumor promoter? [editorial]. *Environ Health Perspect* 1999 Aug;107(8):A390–1.
40. Waters MD, Stack HF, Jackson MA. Short-term tests for defining mutagenic carcinogens. Lyon: International Agency for Research on Cancer; 1999. Scientific publications no 146: 499–536.
41. Cunningham AR, Klopman G, Rosenkranz HS. Identification of structural features and associated mechanisms of action for carcinogens in rats. *Mutat Res* 1998;405(1):9–28.
42. King RD, Srinivasan A. Prediction of rodent carcinogenicity bioassays from molecular structure using inductive logic programming. *Environ Health Perspect* 1996;104 suppl 5:1031–40.
43. Bristol WD, Wachsman JT, Greenwell, A. The NIEHS Predictive-Toxicology Evaluation Project: chemcarcinogenicity bioassays. *Environ Health Perspect* 1996;104S(5):1001–10.
44. Gulezian D, Jacobson-Kram D, McCullough CB, Olson H, Recio L, Robinson D, et al. Use of transgenic animals for carcinogenicity testing: considerations and implications for risk assessment. *Toxicol Pathol* 2000;28(3):482–99.
45. Eastin WC, Haseman JK, Mahler JF, Bucher J. The National Toxicology Program evaluation of genetically altered mice as predictive models for identifying carcinogens. *Toxicol Pathol* 1998;26(4):461–73.
46. Dunson DB, Haseman JK, van Birgelen APJM, Stasiewicz S, Tennant, RW. Statistical analysis of skin tumor data from Tg.AC mouse bioassays. *Toxicol Sci* 2000;55:293–302.
47. Contrera JF, DeGeorge JJ. *In vivo* transgenic bioassays and assessment of the carcinogenic potential of pharmaceuticals. reviews in environmental health, 1998. *Environ Health Perspect* 1998;106 suppl 1.
48. United States Department of Health and Human Services (US-DHHS). Guidance for industry S1B testing for carcinogenicity of pharmaceuticals. Washington (DC): Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER); 1997.
49. Davies TS, Lynch BS, Monro AM, Munro IC, Nestmann ER. Rodent carcinogenicity tests need be no longer than 18 months: an analysis based on 210 chemicals in the IARC monographs. *Food Chem Toxicol* 2000;38(2–3):219–35.
50. IARC. Preamble: occupational exposures to mists and vapours from strong inorganic acids; and other industrial chemicals. Lyon: International Agency for Research on Cancer; 1992. IARC monographs on the evaluation of carcinogenic risks to humans, vol 54:13–32.
51. National Research Council (NRC). Science and judgment in risk assessment. Washington (DC): National Academy Press; 1994.
52. National Toxicology Program (NTP). Introduction: report on carcinogens. 8th edition. Research Triangle Park (NC): National Institute of Environmental Health Sciences; 1998.
53. Sonich-Mullin C, Fielder R, Wiltse J, Baetcke K, Dempsey J, Fenner-Crisp P, et al. IPCS conceptual framework for evaluation a mode of action for chemical carcinogenesis. *Regul Toxicol Pharmacol* 2001;34:146–52.
54. International Life Sciences Institute. A framework for evaluating the human relevance of mode of action data from animals. In press.
55. Fleck C, Braunlich H. Age dependent differences in stimulation and compensation of renal and biliary transport processes: correlation to the physicochemical properties of the exerted substances. *Arzneimittelforschung* 1990;40(9):1047–53.
56. Pavkov KL, Breckenridge C, Campbell W, Capps T, Stevens JT. Evaluation of oncogenic potential of Beacon. *Toxicologist* 1992;12:97.
57. Russell WMS, Burch RL. Principles of humane experimental technique. Springfield (IL): Charles C Thomas, 1959.
58. Interagency Coordinating Committee on the Validation of Alternative Methods (ICVAM). Validation and regulatory acceptance of toxicological test methods. Research Triangle Park (NC): National Institute of Environmental Health Sciences; 1997. NIH publication no 97–3981.
59. Schwetz B, Gaylor D. Alternative tests: carcinogenesis as an example. *Environ Health Perspect* 1998;106 suppl 2:467–71.
60. Combes RD. Detection of nongenotoxic carcinogens: major barriers to replacement of the rodent bioassay. In: van Zutphen LMF, Balls M, editors. Animal alternatives, welfare and ethics: Proceedings of the Second World Congress on Alternatives and Animals in the Life Sciences. Amsterdam: Elsevier; 1997.
61. Spalding JW, French JE, Stasiewicz S, Furedi-Machacek M, Conner F, Tice RR, et al. Responses of transgenic mouse lines p53 (+/–) and Tg.AC to agents tested in conventional carcinogenicity bioassays. *Toxicol Sci* 2000;53:213–23.