Environmental Health Perspectives
Vol. 81, pp. 211-219, 1989

Interspecies Extrapolation in Carcinogenesis: Prediction Between Rats and Mice

by Lois Swirsky Gold,* Leslie Bernstein,† Renae Magaw,* and Thomas H. Slone*

Interspecies extrapolation in carcinogenesis is studied by evaluating prediction from rats to mice and from mice to rats. The Carcinogenic Potency Database, which includes 3500 cancer tests conducted in rats or mice on 955 compounds, is used for the analysis. About half of the chemicals tested for carcinogenicity are positive in at least one test, and this proportion is similar when rats and mice are considered separately. For 392 chemicals tested in both species, 76% of the rat carcinogens are positive in the mouse, and 70% of mouse carcinogens are positive in the rat. When compounds composed solely of chlorine, carbon, hydrogen, and, optionally, oxygen are excluded from the analysis, 75% of mouse carcinogens are positive in the rat. Overall concordance (the percentage positive in both species plus the percentage negative in both) is 76%. Three factors that affect prediction between rats and mice are discussed: chemical class, mutagenicity in the Salmonella assay, and the dose level at which a chemical is toxic. Prediction is more accurate for mutagens than nonmutagens and for substances that are toxic at low (versus only at high) doses. Species differences are not the result of failure in the bioassay to attain the maximum tolerated dose in the negative species or of more frequent testing in the positive species. An analysis of the predictive value of positivity for the 10 most common target sites indicates that most sites are good predictors of carcinogenicity at some site in the other species; the poorest predictors among these common sites are the rat urinary bladder and the mouse liver.

Introduction

Epidemiologic data on chemically induced cancer in humans are difficult to obtain, and only 50 chemicals and manufacturing processes have been evaluated by the International Agency for Research on Cancer (IARC) as having sufficient evidence of carcinogenicity in humans (*1*). In the absence of evidence in humans, chronic-exposure experiments are conducted in rodents, and positive results are used to predict the chemicals that may present a cancer risk to humans. This paper examines extrapolation between species by determining how well one can predict carcinogenicity from a rat to a mouse or from a mouse to a rat. If the carcinogenic response in these two closely related species does not agree, then confidence in extrapolation from rodents to humans (two very different species) may be weakened; conversely, if there is good agreement between rodent species, then confidence may be strengthened.

In a study of interspecies extrapolation in carcinogenesis, results can be presented qualitatively (positive for carcinogenicity vs. not positive), as well as quantitatively, using the carcinogenic potency of the compound (the dose-rate that induces tumors in a predetermined proportion of experimental animals). In an earlier paper (*2*), we discussed the good correlation of carcinogenic potency between rats and mice and showed that the interpretation of this correlation is made difficult by an artifact of potency estimation. In this paper we investigate the qualitative response in rats vs. mice using the largest available database of the results of chronic-exposure rodent experiments, the Carcinogenic Potency Database [CPDB] (*3-5*).

While it is possible to assess the sensitivity of rodent bioassays to detect human carcinogens, it is not possible to assess directly whether rodent carcinogens (of which there are hundreds) have any substantial carcinogenic effect on humans. Many researchers are skeptical of extrapolating risk from rodents to humans (*6,7*). To provide clues about species extrapolation using rats and mice, we

*Biology and Medicine Division, Lawrence Berkeley Laboratory, Berkeley, CA 94720.
†Department of Preventive Medicine, University of Southern California School of Medicine, Los Angeles, CA 90033.
Address reprint requests to L. S. Gold, Biology and Medicine Division, Lawrence Berkeley Laboratory, Berkeley, CA 94720.

first describe the proportion of test agents that have been shown to induce tumors in each species and the relative frequency of each target site. Second, we assess the predictive values of carcinogenicity in one species for results in the other, and we consider three factors that might affect species extrapolation: chemical class, mutagenicity of the compound, and dose level at which the chemical is toxic. Third, we determine whether some target organs are better predictors of a carcinogenic effect in the other species than are other target organs.

Methods

Our analyses are based on the chemicals reported in the CPDB (*3–5*), which is fully described in Gold et al. (*3*). The database includes results for 3500 rat or mouse experiments on 955 compounds obtained from papers published in the general literature through 1984 and from the National Cancer Institute/National Toxicology Program (NCI/NTP) Technical Reports published through May 1986. All experiments in the database meet a specific set of inclusion criteria that are designed to permit the estimation of carcinogenic potency; therefore, reasonable consistency of experimental protocols is assured. Rodent bioassays are included in the database only if the test agent was administered alone, rather than in combination with other substances; if the bioassay included a control group; if the route of administration was either diet, water, gavage, inhalation, IV injection or IP injection; and if the length of experiment was at least 1 year with dosing for at least 6 months. For the CPDB, we do not evaluate the evidence for carcinogenicity in an experiment; rather, we report the evaluation of the published author and calculate the statistical significance of the tumorigenic dose response.

The CPDB includes data on 64% of the 224 substances that have been evaluated by IARC has having sufficient evidence for carcinogenicity in experimental animals (*1*). Some agents are excluded from the CPDB because of route of administration (some polycyclic aromatic hydrocarbons and inorganic chemicals) and some because they are chemical mixtures or industrial processes.

In the following analyses, we classify the results of an experiment as either positive or negative on the basis of the author's opinion in the published paper. In some cases authors do not clearly state their evaluation, and in some NCI/NTP Technical Reports the evidence for carcinogenicity is considered only suggestive; in our analyses we consider these experiments as lacking "clear evidence of carcinogenicity" and classify them as negative. We use the author's opinion to determine positivity because, in addition to statistical significance, it often takes into account historical control rates for particular sites, survival and latency, and/or dose response. Generally, the designation by author's opinion corresponds well with the results of statistical tests for the significance of the dose-response effect (*3*). A two-tailed significance level of 0.01 produces the greatest consistency between author's opinion and statistical significance of the dose-response effect, suggesting that researchers generally follow stringent statistical criteria for determining carcinogenicity.

Of the 3485 experiments in rats or mice in the CPDB, only 5% are evaluated as positive by author's opinion without a corresponding statistically significant result ($p \geq 0.01$), and 6% are not evaluated as positive by the author although the results for at least one target site are statistically significant ($p < 0.01$). These percentages are similar for rats and mice, and for NCI/NTP and literature tests. Of the experiments evaluated as positive by author's opinion that fail to satisfy $p < 0.01$ criterion, 59% satisfy a $p < 0.05$ criterion.

When comparing rats and mice, we report results by chemical; however, the number of experiments per chemical in the CPDB varies. The percentages of chemicals in the CPDB with one experiment, two experiments and more than two experiments are 31, 51, and 18, respectively, for rats and 13, 54, and 33, respectively, for mice. A chemical is considered positive in a given species if it is positive in any single experiment in that species. Positive target sites are identified across experiments in a species using all results for a chemical. Evaluations of the histopathology associated with each target site are not presented because descriptions of tumor pathology vary markedly over time and from paper to paper.

Results

Positivity and Target Sites for 955 Chemicals Tested in Rats or Mice

Table 1 summarizes the proportions of chemicals positive by species and data source. Overall, approximately one-half of the chemicals in the CPDB that were tested in either rats or mice are positive in at least one experiment. Mice appear to be slightly less sensitive than rats in literature reports, but the reverse is true for bioassays of the NCI/NTP. The proportions of positive chemicals for the NCI/NTP are somewhat lower than the general literature, largely because we have classified as negative those NCI/NTP experiments having only suggestive evidence

Table 1. Proportion of chemicals tested in rats or mice that are classified as positive[a] in at least one experiment, by species and reference source, Carcinogenic Potency Database.

Reference source	Proportion carcinogenic in rats or mice (%)	Proportion carcinogenic in rats (%)	Proportion carcinogenic in mice (%)
NCI/NTP or literature[b]	489/955 (51%)	341/706 (48%)	278/645 (43%)
NCI/NTP	120/251 (48%)	84/245 (34%)	94/246 (38%)
Literature	390/771 (51%)	268/502 (53%)	195/444 (44%)

[a]A chemical is classified as positive if the author of at least one published experiment has evaluated the compound as carcinogenic in that species.

[b]Number of chemicals in the NCI/NTP or literature does not equal the sum of each source separately because some chemicals have been reported by both sources.

of carcinogenicity. When positivity is determined by the statistical significance of the dose response (at least one target site with $p < 0.01$) rather than by author's opinion, there is greater consistency between the NCI/NTP and the literature in the proportion of chemicals positive in each species (for rats, 42% NCI/NTP chemicals vs. 49% literature chemicals positive; for mice 44% vs. 45% positive). This results from the fact that the NCI/NTP Technical Reports more often evaluate results at the 0.01 level of statistical significance as lacking "clear evidence of carcinogenicity," whereas authors of general literature papers more often evaluate results for rats at the 0.05 level as positive.

Table 2 reports positive target sites for rat and mouse carcinogens. Twenty-seven sites are identified as positive target organs in the mouse and 30 in the rat. In both species, the liver is the most common target site. The second most common sites are the mouse lung and the rat mammary gland. In the subset of NCI/NTP bioassays, which use an extensive standardized pathology protocol, the most frequent sites are the same for each species, although the rat urinary bladder is identified as frequently as is the mammary gland.

Comparison of Positivity in Rats and Mice for 392 Chemicals in Both Species

The carcinogenic response in rats and mice for all chemicals, by chemical class, and by mutagenicity is summarized in Table 3. For the 392 chemicals, agreement (concordance) in response between rats and mice is 76% (296/392), i.e., 130 chemicals positive in both and 166 negative in both. Based on positivity rates in the entire database (Table 1), we would expect by chance alone that 43% of the chemicals that are carcinogenic in rats would be carcinogenic in mice, and conversely, that 48% of those carcinogenic in mice would be carcinogenic in rats. The positive predictive values for each species among the 392 chemicals (Table 3) are significantly better than would be expected due to chance (chi-square goodness-of-fit test $p < 0.0001$): 76% (130/170) of the rat carcinogens are positive in mice, and 70% (130/186) of the mouse carcinogens are positive in rats. The prediction of negatives is also significantly better than expected: 75% (166/222) for rats and 81% (166/206) for mice (chi-square goodness-of-fit test $p < 0.0001$).

We have examined prediction between species by chem-

Table 2. Frequency of positive target sites for chemicals classified as positive by author's opinion.

Target organ	Number positive at site (%)[a]	
	Chemicals evaluated as carcinogenic in rats (n = 341)	Chemicals evaluated as carcinogenic in mice (n = 278)
Liver	128 (38%)	157 (56%)
Lung	27 (8%)	82 (29%)
Mammary gland	69 (20%)	13 (5%)
Vascular system	22 (6%)	46 (17%)
Stomach	53 (16%)	33 (12%)
Hematopoietic system	34 (10%)	38 (14%)
Kidney/ureter	39 (11%)	11 (4%)
Urinary bladder/urethra	34 (10%)	11 (4%)
Esophagus	27 (8%)	6 (2%)
Nasal cavity/nasal turbinates	26 (8%)	3 (1%)
Ear/Zymbal gland	26 (8%)	2
Skin/subcutaneous tissue	24 (7%)	2
Thyroid gland	20 (6%)	10 (4%)
Small intestine	19 (6%)	3 (1%)
Peritoneal cavity	16 (5%)	7 (3%)
Central nervous system	13 (4%)	2
Oral cavity	13 (4%)	1
Uterus	11 (3%)	5 (2%)
Large intestine	11 (3%)	0
Pituitary gland	6 (2%)	4 (1%)
Adrenal gland	6 (2%)	4 (1%)
Clitoral gland	6 (2%)	2
Preputial gland	1	6 (2%)
Pancreas	6 (2%)	0
Spleen	6 (2%)	0
Harderian gland	0	5 (2%)
Gall bladder	0	3 (1%)
Ovary	0	3 (1%)
Testes	3	1
Myocardium	0	2
Bone	2	0
Mesovarium	2	0
Prostate	2	0
Vagina	1	0

[a]Percentages are not given when fewer than 1% of the carcinogens were active at a given site.

Table 3. Comparison of carcinogenic response in rats and mice, by chemical class.[a]

Chemical class[b] (*n*)	R+M+ (*a*)	R+M− (*b*)	R−M+ (*c*)	R−M− (*d*)	Proportion of R+ that are also M+ [a/(a+b)]	Proportion of M+ that are also R+ [a/(a+c)]
All chemicals (392)	130	40	56	166	76%	70%
Aromatic amines (65)	30	5	14	16	86%	68%
Halogenated compounds						
Chlorinated compounds[c] (50)	18	0	19	13	100%	49%
Other halogenated compounds (23)	13	1	1	8	93%	93%
Miscellaneous aromatics and aliphatics (47)	11	9	4	23	55%	73%
Miscellaneous carbamates and ureas (37)	5	7	2	23	42%	71%
Miscellaneous heterocycles (36)	12	2	3	19	86%	80%
Nitro aromatics and heterocycles (34)	15	2	9	8	88%	63%
Miscellaneous esters and epoxides (31)	6	2	3	20	75%	67%
Azo compounds (18)	5	5	0	8	50%	100%
Inorganic substances (17)	2	1	0	14	67%	100%
Miscellaneous nitrogen compounds, hydrazines, etc. (17)	7	5	1	4	58%	88%
Mixtures or unidentified structures (10)	0	0	0	10		
Nitroso compounds (7)	6	1	0	0	86%	100%
Salmonella results						
Mutagens (138)	64	19	16	39	77%	80%
Nonmutagens (156)	34	16	29	77	68%	54%

[a]Among the 392 chemicals tested in both rats (R) and mice (M), 177 were reported only by NCI/NTP, 150 were reported only in the literature, and 65 were reported by both sources. Four chemicals reported in Table 1 that were tested in both species are not included here because negative results in one species were based solely on a histological exam restricted to only 1 or 2 tissues.

[b]*Chemical classes are ordered by the total number of chemicals in the class. Each chemical is reported in only one class.*

[c]Compounds composed solely of chlorine, carbon, hydrogen, and optionally, oxygen.

ical class (Table 3), but small numbers within most classes prevent detailed, class-specific analyses. Results are most striking for compounds composed solely of chlorine, carbon, hydrogen, and, optionally, oxygen (hereafter, chlorinated compounds): 37/50 (74%) are positive; the predictive value of positivity in rats for positivity in mice is 100% (18/18), compared to only 49% (18/37) from mice to rats. When these chlorinated compounds are eliminated from the set of chemicals tested in both species, the predictive value of positivity from mice to rats is as accurate as that from rats to mice (75%).

Prediction between species is also evaluated on the basis of the mutagenicity of the chemical (Table 3). We obtained evaluations of mutagenicity in Salmonella from recent compilations for 294 of the 392 chemicals in the CPDB that have been tested in both rats and mice (*8,9*; E. Zeiger, personal communication, 1987). A chemical is classified as mutagenic if it is evaluated as positive in any of the sources, and equivocal evaluations are considered as negative. The overall concordance in carcinogenic response between rats and mice for this subset of 294 chemicals is comparable to the larger set of 392 substances. Table 3 indicates that a greater proportion of mutagens are carcinogenic than nonmutagens (72% vs. 51%, chi-square $p < 0.0001$); additionally, a large proportion of carcinogens is not mutagenic (79/178, 44%). A similar result has been reported elsewhere (*9*).

We would expect that chemicals capable of damaging DNA might be more consistently positive across species than chemicals acting by some other mechanism of carcinogenesis. The results in Table 3 provide some support for this hypothesis. Prediction from mouse to rat is significantly more accurate for mutagens (64/80, 80%) than nonmutagens (34/63, 54%, chi-square $p = 0.001$). However, for prediction from rat to mouse the difference is smaller and not significant: 64/83 (77%) for mutagens vs. 34/50 (68%) for nonmutagens (chi-square $p = 0.248$). Because chlorinated carcinogens are all positive in the mouse and

Table 4. Predictive value of carcinogenicity in one species for carcinogenicity in the other, by high dose administered in a bioassay.[a]

Rat carcinogens, high dose in rats, mg/kg/day	Proportion of rat carcinogens that are positive in mice,* (%)	Mouse carcinogens, high dose in mice, mg/kg/day	Proportion of mouse carcinogens that are positive in rats† (%)
< 1	15/16 (94%)	< 1	8/8 (100%)
1-10	23/25 (92%)	1-10	22/27 (81%)
10-100	55/72 (76%)	10-100	35/48 (73%)
100-1000	33/45 (73%)	100-1000	46/71 (65%)
> 1000	4/12 (33%)	> 1000	19/32 (59%)
Total	130/170 (76%)		130/186 (70%)

[a]High dose administered in a positive experiment expressed as the daily dose rate in milligrams/kilogram averaged over the duration of the test. When the CPDB contained more than one positive experiment for a chemical in a species, the high dose rate chosen was from the experiment with the most significant dose-response effect.

*$p = 0.0003$, test for trend.

†$p = 0.008$, test for trend.

are not detected as mutagens, they inflate the predictive value from rat to mouse for nonmutagens. Excluding these chlorinated carcinogens, there is a greater difference between mutagens and nonmutagens in prediction from rats to mice: for mutagens 61/80 (76%) and for nonmutagens 23/39 (59%), (chi-square $p = 0.052$).

A further question related to species extrapolation in carcinogenicity is whether chemicals that are toxic at lower doses are more accurate predictors of carcinogenicity in the second species than chemicals that are toxic only at higher doses. Since, in the standard chronic bioassay, test animals generally are administered a dose that approximates the maximum tolerated dose (MTD), the high dose in a bioassay may be used as a surrogate for the chronic toxicity level of each chemical. Table 4 reports the proportion of rat carcinogens that are positive in the mouse by the magnitude of the high dose administered to rats; parallel data are given for mouse carcinogens.

There is a significant association between the high dose level of a carcinogen used in one species and the likelihood that it produces a carcinogenic response in the other species. The dose-related trend (*10*) is statistically significant both for prediction from rats to mice ($n = 170$, Mantel test for trend, $p = 0.003$) and from mice to rats ($n = 186$, Mantel test for trend, $p = 0.008$). Thus, carcinogens that are toxic at lower doses in a species are better predictors of carcinogenicity in the other species than are less toxic carcinogens.

A possible confounding factor in this significant trend is mutagenicity: since predictive values are more accurate for mutagens than nonmutagens (Table 3), if carcinogens that are toxic at low doses are also more likely to be mutagens, the association between toxic level and predictive value may be a reflection of mutagenicity. We examined the association between mutagenic response and administered dose level in positive rodent tests and found that more toxic carcinogens are significantly more likely to be mutagenic than less toxic carcinogens. Even after adjusting for mutagenicity, however, the dose-related trend in predictive value remains statistically significant from rats to mice ($n = 133$, $p = 0.008$). The adjusted trend predicting from mice to rats is statistically significant when chlorinated compounds are excluded from the set of chemicals used ($n = 110$, $p = 0.007$), but is not statistically significant when all compounds are considered ($n = 143$, $p = 0.116$).

Predictive Value of 10 Frequent Target Sites

To compare extrapolation based on various target organs, we have examined the predictive value of individual target sites in one species for positivity at the same or another site in the second species. Tomatis et al. (*11*) earlier investigated the predictive value of the mouse liver for positivity in the rat using 51 carcinogens, and Ward et al. (*12*) conducted a similar analysis for 85 mouse liver carcinogens. For the 226 positive chemicals in our database tested in both species, there are 10 sites that have been evaluated as target organs for more than 15 chemicals in either the rat or the mouse. In Table 5 we report the number of compounds positive at each of these sites for rats and mice; for each site, we also give the number and proportion that are positive for any site in the second species and the number of chemicals that are positive at the same site in both species. Most sites are good predictors of carcinogenicity at some site in the other species. The least accurate predictors are the urinary bladder in the rat (46%) and the liver in the mouse (63%). For each species, about half the carcinogens are positive in at least one of the same sites in the other species (87/170 for rats, 87/186 for mice).

Many chemicals cause tumors at more than one site in a species. Multiple-site carcinogens in mice are more often positive in rats than single-site carcinogens; this is not the case when predicting from rats to mice. Eighty-two percent (64/78) of the multiple-site mouse carcinogens are positive in the rat compared to 61% (66/108) of the single-site mouse carcinogens (chi-square $p = 0.002$). For rats,

Table 5. Predictive value of target sites in one species for carcinogenicity in a second species: rats and mice.[a]

	Rats		Mice		
Target site[b]	Total number positive at site[c]	Number positive at site in rats also positive at some site in mice (%)	Total number positive at site[c]	Number positive at site in mice also positive at some site in rats (%)	Number positive in same site in rat and mouse
Liver	71	64 (90%)	117	74 (63%)	50
Lung	13	11 (85%)	42	35 (83%)	6
Mammary gland	33	30 (91%)	8	8 (100%)	6
Hematopoietic system	17	12 (71%)	29	24 (83%)	9
Vascular system	12	10 (83%)	26	19 (73%)	5
Urinary bladder/urethra	24	11 (46%)	8	8 (100%)	4
Stomach	23	20 (87%)	19	17 (89%)	9
Kidney/ureter	23	17 (74%)	10	9 (90%)	7
Skin/subcutaneous tissue	18	15 (83%)	2	2 (100%)	1
Ear/Zymbal gland	16	16 (100%)	2	2 (100%)	2
At least one site	170	130 (76%)	186	130 (70%)	87

[a]For chemicals that are tested in both rats and mice and evaluated as positive in at least one experiment.
[b]Target sites are those affected by more than 15 chemicals in at least one species.
[c]Numbers add to more than total for "at least one site" because there is often more than one target site per chemical per species.

78% (68/87) of the multiple-site carcinogens are positive in the mouse compared to 75% (62/83) of the single site carcinogens (chi-square $p = 0.595$). For the poorest predictive organs in Table 5, the rat bladder and the mouse liver, prediction between species is more accurate when the chemical induces tumors at another site as well.

Among compounds that are tested in both rats and mice, it is not common for a positive result to occur in only one target organ of one species. For only three sites are more than 10% of the carcinogens positive in just one site of one species [rat urinary bladder 38% (9/24), mouse liver 26% (31/117) and mouse hematopoietic system 14% (4/29)].

Liver. The liver is the most frequent site in rats and mice as well as the most frequent site in common between them. Positivity in rat liver is highly predictive of positivity in the mouse liver (64/71, 90%), but the mouse liver is more weakly predictive of positivity in the rat (74/117, 63%). The lower predictive value of mouse liver for rat carcinogenicity reflects in part the lower predictive value of chlorinated chemicals as a class, and the fact chlorinated chemicals tend to induce liver tumors. There is a significant difference between chlorinated chemicals and all other chemicals in the predictive value of the mouse liver: among chlorinated compounds 45% (14/31) of those positive in mouse liver are positive in the rat; among other compounds 70% (60/86) of those positive in mouse liver are positive in the rat (chi-square $p = 0.015$). (In the following discussion, we will refer to chemicals that are positive in the mouse liver but not in any other mouse organ or in the rat as "single-site mouse liver carcinogens.")

The frequency of single-site mouse liver carcinogens and the fact that liver tumors occur spontaneously at high rates in the B6C3F1 male has led to the suggestion that carcinogenicity in the mouse liver may more likely be due to promotion of initiated cells or spontaneous tumors than to genotoxicity (*9,12*). We therefore examined whether single-site mouse liver carcinogens are less often mutagenic than other mouse liver carcinogens. For the 91 mouse liver carcinogens on which we have mutagenicity data in Salmonella, the proportion that are mutagenic is lower among single-site mouse liver carcinogens (32%, 8/25) than other mouse liver carcinogens (56%, 37/66 chi-square $p = 0.040$). Importantly, however, there is no difference in the frequency of mutagenicity when chlorinated chemicals are analyzed separately from other chemicals: among the chlorinated compounds, 1/13 (8%) single-site mouse liver carcinogens are mutagenic compared to 2/16 (13%) of other mouse-liver carcinogens (chi-square $p = 0.672$). Among all other compounds 7/12 (58%) single-site mouse liver carcinogens are mutagenic compared to 35/50 (70%) other mouse liver carcinogens (chi-square $p = 0.438$). With respect to the B6C3F1 mouse, we have found that positive mouse liver results in the NCI/NTP bioassays are similar to those of other strains. Among 10 NCI/NTP chemicals that are positive in the B6C3F1 liver and tested in another strain, 7 are positive in the second strain as well. Among the 20 NCI/NTP chemicals that are single-site mouse liver carcinogens, we find no evidence that results are deviant for the male, which has a higher and more variable spontaneous rate of hepatic tumors than the female: 13 of the 20 chemicals (65%) induce liver tumors in both sexes of B6C3F1 mouse, 5 in the male only (25%) and 2 in the female only (10%).

Possible Artifacts in Species Differences of Positivity

Species differences in positivity may reflect artifacts of measurement. For the chemicals tested in NCI/NTP bioassays that are positive in one species only (rat, $n = 23$; mouse, $n = 35$), we examined whether the MTD was administered to the positive species more often than to the negative, since a carcinogenic effect is less likely to be detected at lower doses. We found that the MTD, defined by significant weight-gain depression or the presence of clinical signs of toxicity as cited in the Technical Reports, was usually attained and was obtained equally often in the positive and the negative species.

We also considered the possibility that chemicals positive only in one species may have been tested more frequently in the positive than in the negative species. Among the 96 chemicals tested in both species but positive in only one (Table 3), half were tested equally often in the positive and negative species, and 10% were tested more often in the negative species. Thus, attainment of the MTD and frequency of testing do not account for positive results in one species only.

Carcinogen Identification on the Basis of Two vs. Four Sex-Species Group

The CPDB contains 159 chemicals that have been tested in both sexes of rats and mice and have been evaluated as positive in at least one experiment. This large number permits us to address the question of how many carcinogens would have been identified by performing tests in only two of the four sex-species groups. Table 6 reports the proportion of carcinogens identified by all pos-

Table 6. Predictive value of two sex-species groups for carcinogens tested in both sexes of rats and mice.[a]

	All experiments	NCI/NTP experiments
Sex-species groups used to identify carcinogens[b]	Number identified as carcinogenic at least once, $n = 159$[c]	Number identified as carcinogenic at least once, $n = 114$[c]
MM, MR	146 (92%)	104 (91%)
FM, MR	143 (90%)	102 (89%)
MM, FR	141 (89%)	98 (86%)
FM, FR	139 (87%)	97 (85%)
FM, MM	131 (82%)	92 (81%)
FR, MR	115 (72%)	79 (69%)

[a]For chemicals tested in both sexes of rats and mice that were evaluated as carcinogenic in at least one experiment.

[b]FM, female mice; MM, male mice; FR, female rats; MR, male rats.

[c]Percent indicates proportion that would be correctly identified as carcinogens using results only from experiments in the two sex-species groups, considering as positive an evaluation of carcinogenicity in either sex-species group.

sible combinations of two sex-species groups for the 159 carcinogens. The findings are similar to those reported elsewhere for NCI/NTP bioassays (*13,14*). Testing chemicals in one sex of each species would have resulted in identification of 85 to 92% of the carcinogens that are identified by using four sex-species groups. The greatest number would have been identified on the basis of the male mouse and the male rat. Prediction of positivity between the sexes within each species (80–85%) is better than prediction between species (66–76%). Thus, any combination of a single mouse test and a single rat test would identify more carcinogens than tests in both sexes of either species. One implication of this high detection rate of two groups rather than four is that use of an alternative bioassay design that uses one sex of each species may not result in missing large numbers of rodent carcinogens.

Discussion

Overall Positivity Rate

Our analysis of experiments on 955 chemicals shows that approximately half of the chemicals tested for carcinogenicity in rodents are positive in at least one test, and this proportion is stable across several sets of data (Table 1). The 50% positivity rate agrees with previous results for small groups of chemicals reported by NCI, NTP, and in the general literature (*13–18*). Additionally, for the 294 chemicals that have been tested for both mutagenicity in Salmonella and for carcinogenicity in rats and mice, we find that only 26% are neither carcinogens nor mutagens (Table 3). Positivity rates are so high that it is important to try to understand how representative they might be of the proportion of all chemicals that would be positive if tested in rodent bioassays or the proportion of all chemicals that are potential human carcinogens.

One possible explanation for the high proportion of rodent carcinogens is publication bias, i.e., positive results are more likely to be published than negative results. However, all results are published for NCI/NTP bioassays, and the proportion of positive tests in this group of chemicals is comparable to that of chemicals reported in the literature (Table 1). Another possibility is that more suspicious chemicals are tested for carcinogenicity. For example, chemicals structurally similar to known carcinogens or chemicals that have been shown to be mutagens may be selected more often. This is a likely bias since cancer testing is both expensive and time consuming, and it is prudent to test suspicious compounds. On the other hand, chemicals are selected for testing for several reasons, including the extent of human exposure, level of production, and scientific questions about carcinogenesis. Moreover, while some chemical classes are more often carcinogenic in rodent bioassays than others [e.g., nitroso compounds, aromatic amines, nitroaromatics, chlorinated compounds (Table 3)], predictive capability is still imperfect. For example, most chemicals tested for carcinogenicity are synthetic compounds due to great concern that these man-made chemicals may be carcinogens. However, the positivity rate in the CPDB for naturally occurring chemicals that have been tested in both rats and mice is also high, 31/68 (46%).

It has been suggested that chemicals put on test in the early years of the NCI bioassay program were selected more on the basis of chemical structure, while chemicals tested more recently by NCI or NTP have been selected more on the basis of the extent of human exposure. If this were so, then one might expect the positivity rate to be higher among the earlier chemicals. For the NCI/NTP chemicals we compared positivity rates over time and found no significant difference. Among Technical Reports published before 1979, 51% (60/118) were positive compared to 45% (60/133) published in 1979 or later. Similarly, we found no significant differences in positivity rates over time among chemicals in the general literature (excluding the Innes et al. series discussed below). This suggests that the high positivity rate for chemicals tested in rodents cannot be explained on the basis of these selection criteria.

One large series of mouse experiments conducted by Innes et al. (*19,20*) is often cited as evidence that the proportion of rodent carcinogens is actually low among tested substances (*21*). In this series, among 120 pesticides and industrial chemicals tested in two mouse strains, only 11 (9%) were evaluated as carcinogens. The protocol for this battery of tests differs from that of other tests in the CPDB, and this may account for the lower positivity rate. The Innes series included only 18 animals in the vehicle control group and 18 in a single dose group; the animals were on test for only 18 months, which may not have been an adequate duration to detect a positive effect. (Note: only 16% of the other mouse experiments in the CPDB were of such short duration.) Among the 19 chemicals tested by the Innes group and by another laboratory, the Innes dose was usually lower (sometimes by more than 10 times) than the highest dose administered by others. Because a carcinogenic effect is less likely to be detected at lower doses, these tests were more likely to be negative. Thus, because of experimental design factors, the Innes battery appears to have lacked sensitivity to detect a carcinogenic effect.

Our findings thus suggest that in future testing a high proportion of chemicals may prove to be carcinogens in rodent bioassays. Previously we discussed that carcinogenicity in a rodent bioassay may result from the promotional effects of toxicity when the MTD is administered (*22*). At near-toxic doses, a) cell-killing may remove normal neighboring cells surrounding a mutated cell, thereby removing signals inhibiting growth. This would permit the mutated cell to form a tumor; and b) chronic cell-killing and induced reparative hyperplasia and mutagenic oxidants from phagocytic cells may yield increased gene expression and chance of mutation or associated genetic alteration. Such effects would be expected to increase the proportion of positive chemicals in tests conducted at the MTD, but at low doses without cell killing we would expect the proportion of positives to be lower. The findings that about half of rodent carcinogens are not mutagens suggests that positive bioassay results may be due, in part, to acceleration of the promotional step of carcinogen-

esis. In contrast, most human exposures occur at low doses (compared to the toxic dose in humans). To the extent that positive bioassay results are due to the effects of administering the MTD, the 50% positive rate is probably an overrepresentation of the proportion of all chemicals that are potentially carcinogenic to humans at low doses.

Interspecies Comparisons

We have examined several aspects of interspecies extrapolation in carcinogenesis. When the comparison between rats and mice excludes compounds composed solely of chlorine, carbon, hydrogen, and, optionally, oxygen, the predictive value of positivity in each species for positivity in the other is 75% (i.e., the proportion of carcinogens in one species detected in the other). Positive prediction is slightly less accurate from mouse to rat when these chlorinated compounds are included (70%). These results between two closely related species provide some confidence in the interspecies extrapolation. However, since about half the test agents are carcinogenic in each species, by chance alone we would expect a positive predictive value between species of 50%. The 75% result we obtained would provide greater confidence in interspecies extrapolation if, for example, only 5–10% of the test agents were carcinogens. The overall predictive values we obtained with this large set of chemicals tested in both species are similar to those reported by other researchers for small numbers of test agents (*9,13,14,18,23*). Purchase (*15*) earlier reported a slightly higher value (85%) for a small number of chemicals using different inclusion rules (e.g., route of administration) and different methods to classify positivity. Further discussion of this literature can be found in Haseman and Huff (*14*).

Our analysis has shown that mutagenicity and toxic dose level are both associated with the prediction of carcinogenicity from one rodent species to another. Among carcinogens, positive prediction between species tends to be more accurate for mutagens than for nonmutagens, as well as for chemicals toxic at low doses compared to those that are toxic only at higher doses. Furthermore, there is an association between the toxic level of a chemical and mutagenicity. The explanation for these associations among toxicity, mutagenicity, and predictive value for the second species is not clear. Predictive values between species may be less accurate for nonmutagens than for mutagens because the carcinogenic response for nonmutagens may depend more upon disposition or metabolism, factors that may vary more from species to species than does damage to DNA. The relationship between mutagenicity and toxic dose level may reflect a tendency of mutagens to be more toxic at lower doses because they are effective at killing cells through damaging DNA.

In the standard bioassay, the goal is to administer the test substance at a dose level that is close to the MTD for that substance, regardless of whether the chemical is weak or strong. One possible explanation for the observed association between predictive values and toxicity is that the MTD is less often achieved when the substance is weak, either because it is underestimated or because a higher dose would constitute too great a percentage of the diet. If this were true, then the administration of doses below the MTD could result in fewer positive experiments and less accurate prediction between species for weaker chemicals. We reported a similar finding in an analysis of reproducibility of results in near-replicate experiments where the same chemical was tested more than once by the same route in the same strain, sex, and species (*24*). For these chemicals, the administered dose levels were generally higher for the chemicals that had nonreproducible results than for those that were reproducible.

In our analysis of the predictive value of positivity for the 10 most common target sites in rats and mice, we showed that most sites are good predictors of carcinogenicity at some site in the other species. The mouse liver has been a subject of scientific debate due to its predominance as a target organ and due to its high and variable spontaneous tumor rate in the B6C3F1 hybrid used in the NCI/NTP bioassays. Our result on the predictive value of the mouse liver for positivity in the rat, 63% (74/117), is somewhat lower than the 80% (41/52) reported earlier by Tomatis et al. (*11*), but similar to the 66% (56/85) reported by Ward et al. (*12*). The chemicals and experiments included in the Tomatis et al. data set differ from those in the CPDB; different inclusion rules were used (e.g., Tomatis et al. include experiments conducted by SC injection), and the present analysis includes more chemicals.

Our investigation indicates that the mouse liver is a poorer predictor of rat positivity than are other organs, but that relatively few rodent carcinogens are single-site mouse liver carcinogens (31/226, 14%). Overall, chemicals that induce tumors only in the mouse liver are less often genotoxic than mouse liver carcinogens that are also positive either in another mouse organ or in the rat; however, there is no significant difference in the proportions that are mutagenic when chlorinated chemicals are analyzed separately from nonchlorinated chemicals. Chlorinated compounds are more likely than other carcinogens to be nonmutagenic and positive only in the mouse liver.

Our investigation has described differences between chlorinated compounds and other chemicals with respect to predictivity, mutagenicity, and mouse liver carcinogenicity. A high proportion of these compounds (74%) are carcinogens, and they are among the most visible rodent carcinogens because of their widespread human exposure or high volume production. Many are common solvents or pesticides that are regulated as toxic substances. Further work is needed to understand the mechanism of action of these carcinogens and how this may affect interspecies extrapolation.

The results presented in this paper indicate that prediction of carcinogenicity between rats and mice is correct 75% of the time and that predictive values vary according to several factors including mutagenicity of the compound, chemical class, dose level at which a chemical is toxic, and target organ. In rodent bioassays, even though nearly 50% of the chemicals test positive in each species, overall prediction to the other rodent species is incorrect

25% of the time. Additionally, considering both positive and negative compounds, 25% are discordant between the two rodent species. The practice of extrapolating rodent results to humans should be judged by this value. The overall predictive value from rats to humans or from mice to humans would be expected to be less than 75% since rats and mice are closely related species. To better understand the validity of qualitative extrapolation between rodents and humans, further work is needed on mechanism of action, endogenous damage to DNA, and the extent to which positivity in rodent bioassays may be affected by administering the MTD, as well as on species differences in metabolism, pharmacokinetics, and defenses such as anticarcinogens.

We thank Bruce N. Ames for his continued advice, Jerrold M. Ward, David Freedman, Joseph K. Haseman, and James E. Huff for their suggestions, and Veronica Cabras and Michael Dunker for their assistance. This work was supported by NIEHS/DOE Interagency Agreement 222-Y01-ES-10066 through the Lawrence Berkeley Laboratory.

REFERENCES

1. International Agency for Research on Cancer. Chemicals, Industrial Processes and Industries Associated with Cancer in Humans, IARC Monographs on the Evaluation of Carcinogenic Risk of Chemicals to Humans, IARC Suppl. 7, Lyon, France, 1987.
2. Bernstein, L., Gold, L. S., Ames, B. N., Pike, M. C., and Hoel, D. G. Some tautologous aspects of the comparison of carcinogenic potency in rats and mice. Fundam. Appl. Toxicol. 5: 79–86 (1985).
3. Gold, L. S., Sawyer, C. B., Magaw, R., Backman, G. M., de Veciana, M., Levinson, R., Hooper, N. K., Havender, W. R., Bernstein, L., Peto, R., Pike, M., and Ames, B. N. A Carcinogenic Potency Database of the standardized results of animal bioassays. Environ. Health Perspect. 58: 9–319 (1984).
4. Gold, L. S., de Veciana, M., Backman, G. M., Magaw, R., Lopipero, P., Smith, M., Blumenthal, M., Levinson, R., Gerson, J., Bernstein, L., and Ames, B. N. Chronological supplement to the Carcinogenic Potency Database: Standardized results of animal bioassays published through December 1982. Environ. Health Perspect. 67: 161–200 (1986).
5. Gold, L. S., Slone, T. H., Backman, G. M., Magaw, R., Da Costa, M., Lopipero, P., Blumenthal, M., and Ames, B. N. Second chronological supplement to the Carcinogenic Potency Database: Standardized results of animal bioassays published through December 1984 and by the National Toxicology Program through May 1986. Environ. Health Perspect. 74: 237–329 (1987).
6. Peto, R. Epidemiological reservations about risk assessment. In: Assessment of Risk from Low-Level Exposure to Radiation and Chemicals (A. Woodhead, C. Shellabarger, V. Pond, and A. Hollaender, Eds.), Plenum Press, New York, 1985, pp. 16–30.
7. Freedman, D. A., and Zeisel, H. From mouse to man: The quantitative assessment of cancer risks. Statist. Science 3: 3–56 (1988).
8. Kier, L. E., Brusick, D. J., Auletta, A. E., Von Halle, E. S., Brown, M. M., Simmon, V. F., Dunkel, V., McCann, J., Mortelmans, K., Prival, M., Rao, T. K., and Ray, V. The *Salmonella typhimurium*/mammalian microsomal assay: A report of the U.S. Environmental Protection Agency gene-tox program. Mutat. Res. 168: 69–240 (1986).
9. Zeiger, E. Carcinogenicity of mutagens: Predictive capability of the *Salmonella* mutagenesis assay for rodent carcinogenicity. Cancer Res. 27: 1287–1296 (1987).
10. Breslow, N. E., and Day, N. E., Statistical methods in cancer research. In: The Analysis of Case-Control Studies. IARC Scientific Publication No. 32, Vol. 1, International Agency for Research on Cancer, Lyon, France, 1980, p. 148.
11. Tomatis, L., Partensky, C., and Montesano, R. The predictive value of mouse liver tumour induction in carcinogenicity testing—a literature survey. Int. J. Cancer 12: 1–20 (1973).
12. Ward, J., Griesemer, R. A., and Weisburger, E. K. The mouse liver tumor as an endpoint in carcinogenesis tests. Toxicol. Appl. Pharmacol. 51: 389–397 (1979).
13. Haseman, J. K., Crawford, D. D., Huff, J. E., Boorman, G. A., and McConnell, E. E. Results from 86 two-year carcinogenicity studies conducted by the National Toxicology Program. J. Toxicol. Environ. Health 14: 621–639 (1984).
14. Haseman, J. K., and Huff, J. E. Species correlation in long-term carcinogenicity studies. Cancer Lett. 37: 125–132 (1987).
15. Purchase, I. F. H. Inter-species comparisons of carcinogenicity. Br. J. Cancer 41: 454–468 (1980).
16. Griesemer, R. A. and Cueto, C. Toward a classification scheme for degrees of experimental evidence for the carcinogenicity of chemicals for animals. In: Molecular and Cellular Aspects of Carcinogen Screening Tests (R. Montesano, H. Bartsch, and L. Tomatis, Eds.), IARC Scientific Publication No. 27, International Agency for Research on Cancer, Lyon, France, 1980, pp. 259–281.
17. Chu, K. C., Cueto, C., and Ward, J. Factors in the evaluation of 200 National Cancer Institute carcinogen bioassays. J. Toxicol. Environ. Health 8: 251–280 (1980).
18. DiCarlo, F. J., and Fung, V. A. Summary of carcinogenicity data generated by the National Cancer Institute/National Toxicology Program. Drug Metab. Rev. 15: 1251–1273 (1984).
19. Innes, J. R. M. Evaluation of Carcinogenic, Teratogenic, and Mutagenic Activities of Selected Pesticides and Industrial Chemicals, Vol. 1, Carcinogenic Study. Bionetics Research Laboratories, Inc., National Technical Information Service, Springfield, VA, 1968.
20. Innes, J. R. M., Ulland, B. M., Valerio, M. G., Petrucelli, L., Fishbein, L., Hart, E. R., Pallota, A. J., Bates, R. R., Falk, H. L., Gart, J. J., Klein, M., Mitchell, I., and Peters, J. Bioassay of pesticides and industrial chemicals for tumorigenicity in mice: A preliminary note. J. Natl. Cancer Inst. 42: 1101–1114 (1969).
21. National Cancer Institute. Everything Doesn't Cause Cancer: But How Can We Tell Which Things Cause Cancer and Which Ones Don't? NIH Publication No. 84-2039, revised February, 1984.
22. Ames, B. N., Magaw, R., and Gold, L. S. Ranking possible carcinogenic hazards. Science 236: 271–280 (1987).
23. Haseman, J. K., Huff, J. E., Zeiger, E., and McConnell, E. E. Comparative results of 327 chemical carcinogenicity studies. Environ. Health Perspect. 74: 229–235 (1987).
24. Gold, L. S., Wright, C., Bernstein, L., and de Veciana, M. Reproducibility of results in "near-replicate" carcinogenesis bioassays. J. Natl. Cancer Inst. 78: 1149–1158 (1987).