PREDICTION OF CARCINOGENICITY FROM TWO VERSUS FOUR SEX-SPECIES GROUPS IN THE CARCINOGENIC POTENCY DATABASE

Lois Swirsky Gold, Thomas H. Slone

Life Sciences Division, Lawrence Berkeley Laboratory, Berkeley, California, and Division of Biochemistry and Molecular Biology, Barker Hall, University of California, Berkeley, California

Prediction of a positive result in rodent carcinogenesis bioassays using two instead of four sex-species groups is examined for the subset of chemicals in the Carcinogenic Potency Database that have been tested in four sex-species groups and are positive in at least one (n = 212). Under the conditions of these bioassays, a very high proportion of rodent carcinogens that are identified as positive by tests in four groups is also identified by results from one sex of each species (86–92%). Additionally, chemicals that are classified as "two-species carcinogens" or "multiple-site carcinogens" on the basis of results from four sex-species groups are also identified as two-species or multiple-site carcinogens on the basis of two sex-species groups. Carcinogenic potency (TD50) values for the most potent target site are similar when based on results from two compared to four sex-species groups. Eighty-five percent of the potency values are within a factor of 2 of those obtained from tests in 4 sex-species groups, 94% are within a factor of 4, and 98% are within a factor of 10. This result is expected because carcinogenic potency values are constrained to a narrow range about the maximum dose tested in a bioassay, and the maximum doses administered to rats and mice are highly correlated and similar in dose level.

Information that can be known in advance of a 2-yr bioassay (mutagenicity, class, route, and maximum dose to test) does not identify groups of rodent carcinogens for which four sex-species groups are required to identify carcinogenicity. The range of accurate prediction of carcinogenicity using only male rats and female mice is 93% among mutagens and 88% among nonmutagens; for various routes of administration, 88–100%; for various chemical classes, 75–100%; and for various levels of the maximum dose tested, 81–100%. Results are similar for the pair male rats and male mice.

We are grateful for the suggestions of D. Lai, N. Manley, E. E. McConnell, B. Stern, and J. Swenberg. We thank P. Farvid and K. Chow for technical assistance. This work was supported through the Lawrence Berkeley Laboratory by U.S. Department of Energy contract DE-AC-03-76SF00098, and through the University of California, Berkeley, by National Institute of Environmental Health Sciences Center grant ESO1896.

Journal of Toxicology and Environmental Health, 39:143–157, 1993
0098-4108/93 $10.00 + .00

Using a strength of evidence approach, weaker carcinogens are somewhat less likely than stronger carcinogens to be identified by two sex-species groups. Strength of evidence is measured using the proportion of experiments on a chemical that are positive, the extent to which tumors occur in animals that die before terminal sacrifice, and whether the chemical induces tumors at more than one site and in more than one species.

INTRODUCTION

In an earlier analysis of interspecies extrapolation in carcinogenesis (Gold et al., 1989a), we reported that under the conditions of long-term, chronic bioassays, 85–92% of the chemicals that are identified as carcinogens in rats or mice on the basis of tests in 4 sex-species groups would be identified by testing in 1 sex of each species. Similar findings were reported by Haseman for the bioassays of the National Toxicology Program (Haseman and Huff, 1987). The current standard bioassay design of experiments in male and female rats and male and female mice is expensive. Given the finding that a high proportion of positives would be identified by a two-group design, the Environmental Protection Agency and the National Toxicology Program held a workshop in September 1992 to investigate further the impact of reducing the standard protocol from four to two sex-species groups. This article investigates the impact that a reduced protocol would have on evidence of carcinogenicity among chemicals in our Carcinogenic Potency Database (CPDB) (Gold et al., 1984, 1986, 1987b, 1990, 1993). The number of rodent carcinogens in the CPDB that have been adequately tested in 4 sex-species groups has increased from 159 in our earlier analysis to 212, due to expansion of the CPDB (Gold et al., 1984, 1986, 1987b, 1990, 1993). Thus the impact of a reduced protocol on evaluation of carcinogenicity can now be better assessed, and our analysis here addresses five relevant issues: (1) What proportion of the chemicals identified as rodent carcinogens with a four-group protocol would be identified as carcinogens with a two-group protocol? (2) How much information would be lost because chemicals classified as "two-species carcinogens" or as "multiple-site carcinogens" when tested in four groups, would not be positive in two species or at multiple sites with a two-group design? (3) How would estimates of carcinogenic potency in rodent bioassays be affected? (4) Is it possible to specify, in advance of conducting the bioassays, some types of chemicals for which a reduced protocol would fail to identify most of the rodent carcinogens and that therefore should always be tested in a four-group protocol? (5) Using a strength of evidence approach, are stong carcinogens more likely to be identified by a two-group design than weaker carcinogens?

COMPARISON OF DATA SETS: CHEMICALS TESTED IN FOUR SEX-SPECIES GROUPS VERSUS THE OVERALL CARCINOGENIC POTENCY DATABASE

The CPDB includes the results of all published carcinogenesis bioassays that meet a specified set of inclusion rules, for example, long-term, chronic dosing of the test agent by a route likely to result to whole body exposure (Gold et al., 1984). Positive and negative experimental results that have been published in the general literature or by the National Cancer Institute/National Toxicology Program (NCI/NTP) are included.

In the analysis that follows, a chemical is classified as positive on the basis of the author's opinion in the published paper. Experiments evaluated as "inadequate" by NCI/NTP are excluded. In some cases authors do not clearly state their evaluation, and in some NCI/NTP Technical Reports the evidence for carcinogenicity at a site is considered only "associated" with compound administration or "equivocal"; in our analyses we consider these experiments as lacking positive evidence of carcinogenicity. For NTP reports, the evaluations of "clear" or "some" evidence of carcinogenicity are both classified as positive, as they are by NTP. We use the author's opinion to determine positivity for an experiment because, in addition to statistical significance, it often takes into account historical control rates for particular sites, poor survival, tumor latency, and/or dose response.

To examine the impact of a reduced protocol requires that results be available on a chemical in all four sex-species groups; however, the results of the analysis are intended to be generalized to any chemical that will be tested for regulatory purposes or by the NCI/NTP. Therefore, it is important to know how the data set of chemicals tested in four sex-species differs from the larger group of chemicals tested for carcinogenicity, including those tested in fewer than four groups. Among the 1117 chemicals (positive and negative) in the overall CPDB, only 378 (34%) have been tested in both sexes of rats and mice. Although 75% of these 378 chemicals have been tested by NCI/NTP, only 28% of all chemicals in the overall CPDB have been tested by NCI/NTP. Thus, the four sex-species data set overrepresents NCI/NTP results. Because NCI/NTP bioassays are conducted primarily in Fischer 344 rats and exclusively in $B6C3F_1$ mice, these strains are overrepresented in the four sex-species data set. In the overall CPDB, results from the general literature include 101 mouse strains and 74 rat strains.

By definition, all chemicals in the 4 sex-species data set have been tested in both rats and mice; this compares to 43% in the overall CPDB. Although the chemicals in the 4 sex-species data set average more tests per chemical than the overall CPDB, a similar proportion in the 2 data sets has been shown to be carcinogenic in at least one experiment (52% overall and 56% in the 4 sex-species dataset). Thus, while the positivity rate is similar, in some other respects the four sex-species data set is not representative of the published literature of chronic, long-term carcinogenesis bioassays.

CARCINOGEN IDENTIFICATION ON THE BASIS OF TWO VERSUS FOUR SEX-SPECIES GROUPS

In order to investigate the impact of a reduced protocol, the analysis here is restricted to the 212 chemicals that are positive in at least one experiment when tested in 4 sex-species groups. Table 1 summarizes the results in each sex-species group (including the 166 negative chemicals). A similar proportion of chemicals is positive in each of the 4 sex-species groups (34–39%), indicating that no one group is more sensitive than the others. Since half of the 212 carcinogens are positive in 3 or 4 of the sex-species groups (Table 1), any combination of 2 sex-species groups would have to identify this half of the carcinogens as positive.

The predictive value of positivity for each combination of two sex-species groups is shown in Table 2. Among the 212 chemicals identified as positive on the basis of 4 sex-species groups, a very high proportion (86–92%) is identified as positive using the results from any pair of sex-species groups that consists of one sex of different species (Table 2).

TABLE 1. Summary of Carcinogenicity Results for Studies Evaluated in Male and Female Rats and Mice

Proportion of positive studies	MR	FR	MM	FM	Number of chemicals
4/4	+	+	+	+	71
3/4	+	+	+	–	4
	+	+	–	+	12
	+	–	+	+	10
	–	+	+	+	9
2/4	+	+	–	–	24
	+	–	+	–	2
	+	–	–	+	5
	–	+	+	–	1
	–	+	–	+	3
	–	–	+	+	30
1/4	+	–	–	–	15
	–	+	–	–	6
	–	–	+	–	11
	–	–	–	+	9
0/4	–	–	–	–	166
Proportion positive in each sex-species group	143/378 38%	130/378 34%	138/378 37%	149/378 39%	378

TABLE 2. Predictive Value of Two Sex-Species Groups for CPDB Carcinogens Tested in Both Sexes of Rats and Mice

Sex-species groups used to identify carcinogens	NCI/NTP or literature experiments, number identified as carcinogenic at least once (n = 212)[a]	NCI/NTP experiments, number identified as carcinogenic at least once (n = 149)	Literature experiments, number identified as carcinogenic at least once (n = 59)
MM, MR	194 (92%)	135 (91%)	57 (97%)
FM, MR	194 (92%)	136 (91%)	55 (93%)
MM, FR	183 (86%)	122 (82%)	57 (97%)
FM, FR	184 (87%)	124 (83%)	55 (93%)
FM, MM	167 (79%)	112 (75%)	50 (84%)
FR, MR	162 (76%)	112 (75%)	49 (83%)

[a]For 10 chemicals, data are required from both NCI/NTP and literature. Four chemicals are carcinogenic and tested in 4 sex-species in both NCI/NTP and literature. Two chemicals have ambiguous results because the only positive data are for both sexes combined in the literature (212 = 149 + 59 + 10 − 4 − 2).

Overall, and for the NCI/NTP bioassays separately, the combination of male rats with either female or male mice gives results most similar to those obtained in tests of four sex-species groups, that is, maximizes the identification of carcinogens. In the general literature, the use of male mice and either male or female rats is most accurate, but any combination of 1 sex of each species identifies at least 93% of the positives (Table 2). The proportions identified correctly are slightly higher for the 59 carcinogens from the general literature than from the NCI/NTP (Table 2), and this result provides some confidence in generalizing the finding of accurate identification beyond this data set, despite the overrepresentation of NCI/NTP bioassays.

The fact that a higher proportion of carcinogens is identified by pairs consisting of one sex from each species, rather than by two sexes of the same species, is due to the fact that among chemicals positive in at least one group there is greater agreement (and therefore redundancy of information) between sexes within a species than between species (Table 3).

When evaluating carcinogenicity, additional information can be obtained from rodent bioassays, such as whether the chemical is positive in both rats and mice (two-species carcinogen), or whether the chemical induces tumors in more than one target organ in a species (multiple-site carcinogen). Further confidence in the use of a two-group design would be provided if chemicals that are identified as two-species or multiple-site carcinogens when tested in four sex-species groups were similarly classified using a two-group design. Table 4 reports the accuracy of identification of two-species and multiple-site carcinogens using tests in male rats and female mice or male rats and male mice. A high proportion is identified, and the accuracy for two-species carcinogens

TABLE 3. Proportion of Carcinogens with Redundant Positive Results Between Various Sex-Species Groups

Sex-species group	Proportion positive in both groups among those positive in at least one
MM vs. MR	87/194 (45%)
FM vs. MR	98/194 (51%)
MM vs. FR	85/183 (46%)
FM vs. FR	95/184 (52%)
FM vs. MM	120/167 (72%)
FR vs. MR	111/162 (69%)

is slightly greater for the pair male rat and female mouse than male rat and male mouse.

Carcinogenic potency values (TD_{50}) estimated from studies in either male rats and female mice, or male rats and male mice, are similar to these estimated from four sex-species groups (Table 5). The TD_{50} is defined as the dose to reduce by half the proportion of tumor-free animals at the end of a standard lifetime (Gold et al., 1984; Sawyer et al., 1984). To compare potency values we use the most potent TD_{50} per chemical (Table 5). Among the 194 chemicals accurately identified as positive on the basis of male rats and female mice, the most potent TD_{50} value for 85% of the chemicals is within a factor of 2 of the value obtained from tests in 4 sex-species groups, 94% are within a factor of 4, and 99% are within a factor of 10. Results are similar for the pair male rats and male mice (Table 5). Potency values obtained from different experiments are expected to be similar because given the standard bioassay design, potency is constrained to a narrow range about the maximum dose tested, and the maximum doses administered to rats and mice are similar and highly correlated (Bernstein et al., 1985).

Table 6 reports the 18 chemicals that are classified as positive when tested in four sex-species groups but would be classified as negative if tested only in male rats and female mice. Eight of the 18 are mutagenic

TABLE 4. Accuracy of Identification of Multiple Species and Multiple Site Carcinogenicity Using Two Sex-Species Groups (MR, FM or MR, MM)

Accuracy of identification	Proportion identified by MR, FM (%)	Proportion identified by MR, MM (%)
Proportion identified as positive in two species	98/117 (84%)	87/117 (74%)
Proportion identified as multiple site within a species	118/121 (98%)	118/121 (98%)

TABLE 5. Ratio of Carcinogenic Potency Values (Most Potent TD_{50}) from Experiments in Two Versus Four Sex-Species Groups

	Identified by MR-FM		Identified by MR-MM	
Ratio	*n*	%	*n*	%
<2	165	85	170	87
2–2.99	13	7	9	5
3–3.99	5	3	2	1
4–9.99	8	4	9	5
≥10	3	1	4	2
Total	194	100	194	100

in *Salmonella* and 10 are not mutagenic (Haworth et al., 1983; Kier et al., 1986; Mortelmans et al., 1986; Zeiger, 1987, 1990; Zeiger et al., 1987, 1988, 1992; E. Zeiger, personal communication; A. E. Auletta, personal communication). Fourteen are positive in only a single sex-species or a single species. Table 8 indicates that chemicals positive in only one target site in a single sex-species, thus indicating that they have relatively weak evidence of rodent carcinogenicity (see below).

TABLE 6. Rodent Carcinogens Not Identified by Tests in MR and FM

Mutagens (*n* = 8)	Target tissue
3-Amino-4-ethoxyacetanilide	Thyroid (MM)
2-Amino-5-nitrothiazole	Kidney, lung, mammary (FR)
Nithiazide	Liver (MM), mammary (FR)
3-Nitro-*p*-acetophenetide	Liver (MM)
C.I. acid orange 3	Kidney (FR)
p-Quinone dioxime	Bladder (FR)
Styrene	Mammary (FR)
FD&C Violet no. 1	Mammary, skin (FR)
Nonmutagens (*n* = 10)	**Target tissue**
Butylated hydroxytoluene	Liver, lung (MM)
Chlorinated paraffins	Hematopoietic (MM)
Dicofol	Liver (MM)
Hexanamide	Hematopoietic (MM)
Pentachloronitrobenzene	Liver (MM)
Piperonyl sulfoxide	Liver (MM)
Tetrachlorvinphos	Liver (MM)
p-Tolyurea	Hematopoietic (MM)
Trimethylthiourea	Thyroid (FR)
C.I. Vat Yellow 4	Hematopoietic (MM)

In our data set of 212 rodent carcinogens that have been tested in 4 groups, there are 8 chemicals known to cause cancer in humans (IARC, 1987; Tomatis and Bartsch, 1990). All eight are identified as positive by tests in the pair male rats and female mice or the pair male rats and male mice (aflatoxin B_1, benzene, chlorambucil, cyclophosphamide, diethylstilbestrol, melphalan, thio-tepa, and vinyl chloride).

CAN SOME TYPES OF CHEMICALS BE IDENTIFIED FOR WHICH A TWO-GROUP PROTOCOL WOULD FAIL TO IDENTIFY MOST OF THE RODENT CARCINOGENS?

Some information can be determined prior to conducting a 2-yr bioassay, such as the mutagenicity of a compound (Haworth et al., 1983; Kier et al., 1986; Mortelmans et al., 1986; Zeiger, 1987, 1990; Zeiger et al., 1987, 1988; 1992; E. Zeiger, personal communication; A. E. Auletta, personal communication), its chemical class, route of administration, and maximum tolerated dose (MTD). It is possible that this information might be used to define some groups of chemicals for which a two-group design would fail to identify a large proportion of rodent carcinogens. Our analysis, however, suggests that this is not the case because a high proportion of every group is identified by a reduced protocol (Table 7). Results are similar for the pair male rats and female mice and the pair male rats and male mice. For example, with male rats and female mice, among the chemicals identified as positive by 4 groups, 93% of the mutagens are identified as positive, compared to 88% of the nonmutagens (Table 7). For various chemical classes 75–100% are identified by tests in 2 groups (Table 7); the small number of carcinogens in some classes makes evaluation difficult. No route of administration or level of MTD stands out as failing to identify carcinogens on the basis of two groups (Table 7).

STRENGTH OF EVIDENCE OF CARCINOGENICITY AND IDENTIFICATION OF RODENT CARCINOGENS USING A REDUCED PROTOCOL

If the rodent carcinogens with the strongest evidence for carcinogenicity were less likely than weaker carcinogens to be identified by a two-group protocol, then this would be an undesirable impact of a reduced protocol. We have investigated strength of evidence in several ways and find that the stronger carcinogens are more likely to be identified by a two-group design than the weaker carcinogens. We have analyzed results based on the pair male rats and female mice as well as the pair male rats and male mice. In Table 8 we evaluate strength of evidence using the number of species that are positive and the number of target sites. By necessity, carcinogens that are positive in more sex-

TABLE 7. Proportion of Carcinogens Identified by Two Sex-Species (MR, FM or MR, MM) by Mutagenicity, Chemical Class, Route of Administration, and Toxicity of Chemical

Parameter	Proportion identified by MR, FM (%)		Proportion identified by MR, MM (%)	
Mutagenicity in Salmonella				
Mutagens	103/111	(93%)	99/111	(89%)
Nonmutagens	73/83	(88%)	79/83	(95%)
Chemical class				
Chlorinated compounds	37/39	(95%)	39/39	(100%)
Other halogenated compounds	16/16	(100%)	15/16	(94%)
Aromatic amines	38/40	(95%)	37/40	(93%)
Nitro aromatics and heterocyclics	18/23	(78%)	18/23	(78%)
Miscellaneous aromatics and aliphatics	15/20	(75%)	18/20	(90%)
Miscellaneous esters and epoxides	18/19	(95%)	17/19	(89%)
Miscellaneous nitrogen compounds, hydrazines, etc.	15/16	(94%)	15/16	(94%)
Miscellaneous heterocycles	16/16	(100%)	14/16	(88%)
Miscellaneous carbamates and ureas	6/8	(75%)	6/8	(75%)
Azo compounds	7/7	(100%)	7/7	(100%)
Nitroso compounds	5/5	(100%)	5/5	(100%)
Inorganic substances	3/3	(100%)	3/3	(100%)
Route of administration				
Diet	115/130	(88%)	119/130	(92%)
Gavage	70/73	(96%)	68/73	(93%)
Water	24/24	(100%)	23/24	(96%)
Inhalation	19/20	(95%)	19/20	(95%)
Intraperitoneal injection	14/14	(100%)	12/14	(86%)
Intravenous injection	3/3	(100%)	3/3	(100%)
Highest dose tested				
<1	6/6	(100%)	6/6	(100%)
1–10	13/13	(100%)	12/13	(92%)
10–100	53/55	(96%)	51/55	(93%)
100–1000	88/96	(92%)	87/96	(91%)
≥1000	34/42	(81%)	38/42	(90%)

species groups are more likely to be identified with a two-group protocol than are chemicals positive in only a single sex-species or a single species. Table 8 indicates that chemicals are positive in only one species at a single target site are least likely to be identified by tests in two groups.

Table 9 indicates that when only one-fourth of the individual experiments on a chemical are positive, then the identification of carcinogenicity is less accurate with a reduced protocol (some chemicals are tested more than once in a sex-species group).

Lethality of tumors (or latency) is another measure of strength of carcinogenicity: If tumors occur early in animals that die before terminal sacrifice, then the evidence of carcinogenicity is stronger than if

TABLE 8. Proportion of Carcinogens Identified by Tests in Two Sex-Species by Strength of Evidence for Carcinogenicity

Strength of evidence	Proportion identified by MR-FM	Proportion identified by MR-MM
Multiple site in both rats and mice	43/43 (100%)	43/43 (100%)
Multiple site in one species, single in other	42/42 (100%)	42/42 (100%)
Single site in both species	29/30 (97%)	27/30 (90%)
Multiple site in one species, negative in other	33/36 (92%)	33/36 (92%)
Single site in one species, negative in other	43/57 (75%)	45/57 (79%)

Note. The total number of carcinogens in this table (n = 208) differs from that in Table 2 (n = 212). This difference is due to the fact that multiple-site carcinogenesis cannot be measured for experiments that restrict histopathological examination or report data for only a few selected tissues. The exclusion of such experiments results in a smaller number of chemicals in this table.

tumors are found in sacrificed animals. Table 9 reports the proportion of carcinogens identified with two groups by the extent to which tumors at some site occur in animals that die before terminal sacrifice. This measure, the "sacrifice ratio," is the ratio of the TD_{50} calculated for all animals to the TD_{50} for animals that die before sacrifice. If all of the animals die before sacrifice, then the ratio will be one. If nearly all tumors occur at sacrifice, then the TD_{50} for all animals (numerator) will

TABLE 9. Proportion of Carcinogens Identified by Two Sex-Species (MR,FM or MR,MM) by Positivity Rate and Lethality of Tumors

	Proportion identified by MR, FM (%)	Proportion identified by MR, MM (%)
Proportion of experiments that are positive		
1	46/46 (100%)	46/46 (100%)
0.75–0.99	36/36 (100%)	36/36 (100%)
0.50–0.74	69/70 (99%)	67/70 (96%)
0.25–0.49	10/11 (91%)	10/11 (91%)
<0.25	33/49 (67%)	35/49 (71%)
Extent to which tumors occur in animals dead before terminal sacrifice (sacrifice ratio)		
0.9–1	27/27 (100%)	27/27 (100%)
0.6–0.89	17/18 (94%)	17/18 (94%)
0.3–0.59	50/53 (94%)	47/53 (89%)
<0.3	42/51 (82%)	44/51 (86%)

be lower than the TD_{50} calculated for animals dying before sacrifice (denominator), and the ratio will be low. Table 9 indicates that when tumors are found primarily in animals that died before terminal sacrifice, the chemical is more likely to be identified as carcinogenic with a two-group protocol than when tumors occur primarily at terminal sacrifice.

Taken together, these three findings on strength of evidence of carcinogenicity (multiple species and site carcinogenicity, proportion of experiments that are positive, and lethality of tumors) indicate that the weakest rodent carcinogens are more likely to be missed using a reduced protocol.

DISCUSSION

Our results have shown that under the conditions of chronic bioassays, a high proportion (86–92%) of rodent carcinogens that are identified using a protocol of four sex-species groups can be identified with two groups consisting of one sex from each species. This result is consistent with our earlier analysis (Gold et al., 1989a), as well as with the analysis of Haseman on the NCI/NTP bioassays (Haseman and Huff, 1987). Additionally, we have shown that high proportions of carcinogens can be identified for all chemical classes, of both mutagens and nonmutagens, and for carcinogens with various levels of toxicity (MTD). Estimates of carcinogenic potency are not greatly different when based on experiments in only one sex of each species rather than four sex-species groups. Moreover, the stronger the evidence of carcinogenicity for a chemical based on tests in four groups, the more likely it is to be identified by two groups; a high proportion of weaker carcinogens is also identified by two groups. The rodent carcinogens that are missed do not tend to have strong evidence of carcinogenicity. Thus, the overall results of our analysis suggest that a reduced protocol merits further consideration.

A two-group protocol, however, does not identify 100% of the rodent carcinogens that are identified by four groups. A benchmark for evaluating the importance of this less than perfect identification is provided by our earlier analysis of reproducibility of results in "near-replicate" experiments (Gold et al., 1987a, 1993). We compared results of experiments in which the same chemical had been tested more than once in the same strain and sex of animal by the same route of administration. We concluded that among chemicals positive in one experiment, 21% (18/87) are not positive in the near-replicate experiment. Similarly, potency estimates (TD_{50}) based on near-replicate experiments (Gold et al., 1989b) vary somewhat, as do the potency estimates based on two versus four sex-species groups (see Table 5); using the near-replicate comparisons as a benchmark, the variation in TD_{50} in Table 5

compares favorably. Thus, some discrepancies in results are to be expected between two- and four-group protocols, since reproducibility is not perfect even for near-replicate tests.

If a two-group protocol were to be adopted, which two groups should be selected? We have found that for the chemicals in the CPDB and the subset of chemicals tested by NCI/NTP, the most accurate identification is obtained by using either the pair male rats and female mice or the pair male rats and male mice. Results are similar for these two pairs. However, for the 59 carcinogens from the general literature data set of chemicals tested in four sex-species groups, the proportion identified is higher for all pairs consisting of one sex of each species, and is highest for the pairs male mice and male rats, or male mice and female rats. These literature results are based on a variety of strains of rats and mice. Thus, it is not clear that the most accurate identification would always be achieved by selection of male rats in combination with either male or female mice. In a given future bioassay, it may not be necessary to select in advance which two sex-species groups to use in a particular bioassay. If the 13-wk subchronic study were to be conducted with all 4 sex-species groups, then results on toxicity and histopathology (nonneoplastic lesions) from that study could be used in selecting the sex-species groups likely to be most sensitive in a 2-yr bioassay. Our analysis strongly suggests that one sex of each species would be desirable; which particular two groups to test, however, might be varied depending on results of subchronic toxicity studies. One disadvantage of choosing different sex-species pairs for bioassays of various chemicals is that comparisons of results on different chemicals would be more difficult, since the sex-species groups would vary as well as the chemical agent.

How much confidence should we have in generalizing beyond the chemicals in this analysis the finding that a two-group protocol can identify a high proportion of carcinogens that are identified by a four-group protocol? Our results are based on only 378 of the 1117 chemicals in the CPDB, because only 378 have been tested in 4 sex-species groups. In this subset of the CPDB the NCI/NTP bioassays are overrepresented compared to the overall database: Whereas 28% of the chemicals in the overall CPDB are NCI/NTP chemicals, the proportion in the 4 sex-species data set is 75%, as already discussed. A reduced protocol might conceivably predict positivity less well in a broader data set because strains or chemical classes that are underrepresented in the four sex-species data set might be used. For example, the NCI/NTP bioassays use only $B6C3F_1$ mice and primarily Fischer 344 rats, and results might not be the same for all strains. Our findings in Table 4 provide some confidence in generalizability because identification of carcinogens is slightly more accurate for the general literature than for NCI/NTP bioassays. Additional information about generalizability can be obtained by

examining results in the CPDB for chemicals that have not been tested in four groups but that have been tested in both males and females of species. Sensitivity to detect positive chemicals would not be weakened with a two-group protocol using one sex of each species if agreement in positivity (redundancy) between males and females is as frequent for other chemicals as it is for the four sex-species data set. If redundancy between the sexes of a species is great, then testing in only one sex of each species would likely identify a high proportion of the carcinogens that would be identified by testing four groups. We have examined the positive chemicals in the literature that have not been tested in four groups but have been tested in two sexes of either rats or mice, and we have excluded experiments on Fischer rats and $B6C3F_1$ mice. The redundancy between males and females of each species is as follows: In rats, 79% (53/67) of the carcinogens have positive results in both sexes, and in mice 75% (73/97). The respective proportions in the 4 sex-species data set are 69% (111/162) in rats and 72% (120/167) in mice (Table 5). Thus, redundancy is slightly greater for carcinogens that are not in the four sex-species data set, providing further confidence in the use of a two-group protocol in future bioassays.

It should be possible to obtain further information about the impact of a reduced protocol and the choice of which sex-species groups to use for a particular strain, by analyzing results from laboratories that conduct carcinogenesis bioassays regularly, for example, pharmaceutical laboratories or chemical manufacturers. The EPA database of bioassay results on chemicals submitted for registration is another potential source of such data. Analyses like the one in this article could be conducted for each of those data sets.

With respect to generalizing our results to all chemical classes, we note that our 4 sex-species data set underrepresents nitrosamines and hydrazines when compared to the overall CPDB. We therefore examined results for nitrosamines and hydrazines that have been tested in both sexes of rats or mice in the overall CPDB, in order to assess extent of redundancy between the sexes for these chemical classes. If there is great redundancy, then testing one sex of each species in a two-group protocol should not result in less accurate identification of rodent carcinogens than we have found in our analysis. Experiments of nitrosamines are conducted primarily in rats, and experiments of hydrazines are primarily in mice. Our results indicate that the redundancy in each of these species is somewhat greater between the sexes for these two classes than for the four sex-species data set. Therefore, it is reasonable to expect that a reduced protocol would identify most of the carcinogens in these classes. Thus, our additional analyses of redundancy between results in males and females for strains other than Fischer 344 rats or $B6C3F_1$ mice, and for chemical classes that are underrepresented in the 4 sex-species data set, provide further confidence in generalizing our results beyond the 378 chemicals tested in 4 sex-species groups.

If a reduced protocol were to be adopted, there would be fewer animals, reduced work required for conducting the study and performing histopathological examination, etc. White some of the cost savings might be used to test additional chemicals, consideration should be given to adding studies of mechanisms of carcinogenesis to the standard protocol. The results of carcinogenesis bioassays are routinely used to assess human risk from chemical exposure; however, as currently conducted, animal cancer tests do not provide sufficient information to estimate potential risk to humans (Gold et al., 1992). We have postulated (Ames and Gold, 1990a, 1990b; Ames et al., in press) that chronic administration of chemicals at the maximum tolerated dose (MTD) increases cell division (mitogenesis), which in turn increases rates of mutagenesis and, thus, carcinogenesis. If mitogenesis is a dominant factor in carcinogenesis at the MTD, then at low doses where mitogenesis is not generally induced, the hazards to humans of rodent carcinogens may be much lower than commonly assumed. Adding routine measurements of mitogenesis to the 13-wk toxicology study and the 2-yr bioassay would provide information that would improve dose setting, interpretation of experimental results, and risk assessment.

REFERENCES

Ames, B. N., and Gold, L. S. 1990a. Perspective: Too many rodent carcinogens: Mitogenesis increases mutagenesis. *Science* 249:970-971. [Also see letters in *Science* 250:1498 (1990); 250:1645-1646 (1990); 251:12-13 (1991); 251:607-608 (1991); 252:902 (1991).]

Ames, B. N., and Gold, L. S. 1990b. Chemical carcinogenesis: Too many rodent carcinogens. *Proc. Natl. Acad. Sci. USA* 87:7772-7776.

Ames, B. N., Shigenaga, M. K., and Gold, L. S. In press. DNA lesions, inducible DNA repair, and cell division: Three key factors in mutagenesis and carcinogenesis. *Environ. Health Perspect.*, in press.

Bernstein, L., Gold, L. S., Ames, B. N., Pike, M. C., and Hoel, D. G. 1985. Some tautologous aspects of the comparison of carcinogenic potency in rats and mice. *Fundam. Appl. Toxicol.* 5:79-86.

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. C., and Ames, B. N. 1984. A Carcinogenic Potency Database of the standardized results of animal bioassays. *Environ. Health Perspect.* 58:9-319.

Gold, L. S., de Veciana, M., Backman, G. M., Magaw, R., Lopipero, P., Smith, M., Blumenthal, M., Levinson, R., Bernstein, L., and Ames, B. N. 1986. Chronological supplement to the Carcinogenic Potency Database: Standardized results of animal bioassays published through December 1982. *Environ. Health Perspect.* 67:161-200.

Gold, L. S., Wright, C., Bernstein, L., and de Veciana, M. 1987a. Reproducibility of results in 'near-replicate' carcinogenesis bioassays. *JNCI* 78:1149-1158.

Gold, L. S., Slone, T. H., Backman, G. M., Magaw, R., Da Costa, M., Lopipero, P., Blumenthal, M., and Ames, B. N. 1987b. 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.

Gold, L. S., Bernstein, L., Magaw, R., and Slone, T. H. 1989a. Interspecies extrapolation in carcinogenesis: Prediction between rats and mice. *Environ. Health Perspect.* 81:211–219.

Gold, L. S., Slone, T. H., and Bernstein, L. 1989b. Summary of carcinogenic potency (TD_{50}) and positivity for 492 rodent carcinogens in the Carcinogenic Potency Database. *Environ. Health Perspect.* 79:259–272.

Gold, L. S., Slone, T. H., Backman, G. M., Eisenberg, S., Da Costa, M., Wong, M., Manley, N. B., Rohrbach, L., and Ames, B. N. 1990. Third chronological supplement to the Carcinogenic Potency Database: Standardized results of animal bioassays published through December 1986 and by the National Toxicology Program through June 1987. *Environ. Health Perspect.* 84:215–286.

Gold, L. S., Manley, N. B., Slone, T. H., Garfinkel, G. B., Rohrabach, L., and Ames, B. N. 1993. 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.* 100:65–135.

Gold, L. S., Manley, N. B., and Ames, B. N. 1992. Extrapolation of carcinogenesis between species: Qualitative and quantitative factors. *Risk Anal.,* 12:579–588.

Haseman, J. K., and Huff, J. E. 1987. Species correlation in long-term carcinogenicity studies. *Cancer Lett.* 37:125–132.

Haworth, S., Lawlor, T., Mortelmans, K., Speck, W., and Zeiger, E. 1983. *Salmonella* mutagenicity test results for 250 chemicals. *Environ. Mutagen.* 5 (Suppl. 1):3–142.

International Agency for Research on Cancer. 1987. *Overall Evaluations of Carcinogenicity,* suppl. 7. Lyon, France: IARC.

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. 1986. The *Salmonella typhimurium*/mammalian microsomal assay: A report of the U.S. Environmental Protection Agency gene-tox program. *Mutat. Res.* 168:69–240.

Mortelmans, K., Haworth, S., Lawlor, T., Speck, W., Tainer, B., and Zeiger, E. 1986. *Salmonella* mutagenicity tests: II. Results from the testing of 270 chemicals. *Environ. Mutagen.* 8(suppl. 7):1–119.

Sawyer, C., Peto, R., Bernstein, L., and Pike, M. C. 1984. Calculation of carcinogenic potency from long-term animal carcinogenesis experiments. *Biometrics* 40:27–40.

Tomatis, L., and Bartsch, H. 1990. The contribution of experimental studies to risk assessment of carcinogenic agents in humans. *Exp. Pathol.* 40:251–266.

Zeiger, E. 1987. Carcinogenicity of mutagens: Predictive capability of the *Salmonella* mutagenesis assay for rodent carcinogenicity. *Cancer Res.* 27:1287–1296.

Zeiger, E. 1990. Mutagenicity of 42 chemicals in *Salmonella*. *Environ. Mol. Mutagen.* 16(suppl. 18):32–54.

Zeiger, E., Anderson, B., Haworth, S., Lawlor, T., Mortelmans, K., and Speck, W. 1987. *Salmonella* mutagenicity tests: III. Results from the testing of 255 chemicals. *Environ. Mol. Mutagen.* 9(suppl. 9):1–110.

Zeiger, E., Anderson, B., Haworth, S., Lawlor, T., and Mortelmans, K. 1988. *Salmonella* mutagenicity tests: IV. Results from the testing of 300 chemicals. *Environ. Mol. Mutagen.* 11(suppl. 12):1–158.

Zeiger, E., Anderson, B., Haworth, S., Lawlor, T., and Mortelmans, K. 1992. *Salmonella* mutagenicity tests: V. Results from the testing of 311 chemicals. *Environ. Mol. Mutagen.* 19(suppl. 21):2–141.

Received December 14, 1992
Accepted January 11, 1993