important molecular evidence that supports the concept of multistage carcinogenesis, presented a questionable method (Abbots' formula) to account for background tumor incidence, distorted the IARC definitions of sufficient and limited evidence and failed to adjust for a major confounding factor (age) in their analysis of the DDT data. In spite of all this, I share in large measure their skepticism about the scientific value of routine risk assessments that use statistical models fitted to limited animal data obtained at high doses to predict the human response at low ones. Society needs critics like Freedman and Zeisel to challenge establishment viewpoints, lest the repeated use of "inference guidelines" such as low-dose linear extrapolation lends them undeserved credence. Hopefully, other scientists will continue their constructive efforts to improve the biological and statistical models and to contribute their expertise to the decision making process.

ADDITIONAL REFERENCES

BARRETT, J. C. (ed.) (1987). A multistep model for neoplastic development: Role of genetic and epigenetic changes. In *Mechanisms of Environmental Carcinogenesis II: Multistep Models of Carcinogenesis*. CRC Press, Boca Raton, Fla.

BEIR (1980). *The Effects on Populations of Exposure to Low Levels of Ionizing Radiation*. Report of the Advisory Committee on the Biological Effects of Ionizing Radiations, National Academy of Sciences, Washington.

BRESLOW, N. E. and BECKWITH, J. B. (1982). Epidemiological features of Wilms' tumor: Results of the National Wilms' Tumor Study. *J. Nat. Cancer Inst.* **68** 429–436.

BRESLOW, N. E., ET AL. (1974). Associations between tumor types in a large-scale carcinogenesis study of CF-1 mice. *J. Nat. Cancer Inst.* **52** 233–239.

CALIFORNIA STATE DEPARTMENT OF HEALTH SERVICES (1982). Carcinogen identification of policy: A statement of science as a basic of policy.

DINSE, G. E. and LAGAKOS, S. W. (1983). Regression analysis of tumor prevalence data. *Appl. Statist.* **32** 236–248.

FRANCKE, V., ET AL. (1979). Aniridia-Wilms' tumor association: Evidence for specific deletion of 11p13. *Cytogenet. Cell. Genet.* **24** 185–192.

GILBERT, E. S. (1985). How much can be learned from populations exposed to low levels of radiation? *Statistician* **34** 19–30.

KNUDSON, A. G. and STRONG, L. C. (1972). Mutation and cancer: A model for Wilms' tumor of the kidney. *J. Nat. Cancer Inst.* **48** 313–324.

KOUFOS, A., ET AL. (1984). Loss of alleles at loci on human chromosomes during genesis of Wilms' tumor. *Nature* **309** 170–172.

PETO, R. (1978). Carcinogenic effects of exposure to very low levels of carcinogenic substances. *Environ. Health Perspect.* **22** 155–159.

Comment

J. K. Haseman

I was disappointed to find Freedman and Zeisel taking such a one-sided and negative position concerning the scientific value of laboratory animal studies for assessing possible cancer risks to humans. The scientific merits of laboratory animal studies and quantitative risk estimation have been debated for years, and Freedman and Zeisel raise no new points that have not been considered extensively elsewhere. The difference between their article and more definitive publications (e.g., Office of Science and Technology Policy, 1985) is that Freedman and Zeisel make no effort to present a balanced view on the major issues.

Freedman and Zeisel utilize several questionable techniques to achieve their objectives. These include (1) selectively citing references that appear to support their point of view while ignoring other publications that express contrary views, and (2) misrepresenting or misinterpretating data from various sources, in some cases reaching the opposite conclusion to that given by the original investigator. Examples of this will be given throughout these comments, which will be limited primarily to the area of qualitative risk assessment.

Throughout the paper Freedman and Zeisel display an arrogant attitude toward nonstatisticians, (e.g., assuming that investigators do not randomize properly unless the randomization scheme is stated explicitly; claiming that "pathologists see themselves as professionals exempt from bias"). This air of superiority, especially when considering biological issues, reduces their own credibility and the credibility of all statisticians in the eyes of biologists, many of whom feel that statisticians and lawyers debating science is no more meaningful than biologists debating p-values.

The major criticisms of laboratory animal carcinogenicity studies cited by Freedman and Zeisel include the following.

J. K. Haseman is Research Mathematical Statistician, Biometry and Risk Assessment Program, National Institute of Environmental Health Sciences, P. O. Box 12233, Research Triangle Park, North Carolina 27709.

1. THERE IS LITTLE OR NO SCIENTIFIC BASIS FOR USING LABORATORY ANIMALS IN GENERAL AND RODENTS IN PARTICULAR FOR CARCINOGENICITY TESTING

Freedman and Zeisel are critical of laboratory animal carcinogenicity studies in general and rodent studies in particular. In their view, the only scientific justification for using mice and rats is that "they are small, cheap and easy to maintain under laboratory conditions. Furthermore, experimentalists have much experience with rats and mice. There seem to be no other serious arguments for using these two test species in cancer testing," citing the 1981 Report of the Office of Technology Assessment as the basis for their conclusion. Freedman and Zeisel further conclude that current laboratory animal studies "have little to do with basic research" and do not "contribute much to the scientific regulation of health or environmental hazards."

However, the primary basis for using laboratory animals such as rats and mice is scientific, not pragmatic. Freedman and Zeisel fail to note that the Office of Technology Assessment Report (1981) also concluded that "metabolic studies have shown that most differences between humans and experimental animals are quantitative rather than qualitative and support the idea that animal results can be used to predict human responses" (page 126). The report also quotes (page 122) the National Academy of Sciences' National Research Council, which concluded that "there are large bodies of experimental data that indicate that exposures that are carcinogenic to animals are likely to be carcinogenic to man, and vice versa."

Freedman and Zeisel further state that an "oft-recited argument" to defend current laboratory animal studies is "that humans and mice are both mammalian species," a view they sarcastically refer to as "verges on sentimentality." However, the similarities between humans and laboratory animals are far more important than this.

For example, "clearly the characteristics of chemical carcinogenesis in animals and in humans appear to be identical" (Huff and Moore, 1984) and "the experimental test systems used to evaluate the toxicities of potential anticancer drugs [in laboratory animals] correlate remarkably closely with the results in man" (Freireich et al., 1966). Moreover, "molecular, cellular, tissue and organ functions are strikingly similar in all animal species" (Rall, 1979). In summary, "experimental evidence to date certainly suggest that there are more physiologic, biochemical, and metabolic similarities between laboratory animals and humans than there are differences. These similarities increase the probability that results observed in a laboratory setting will predict similar results for humans. Clearly the accumulated experience in the field of carcinogenesis supports this concept" (Rall et al., 1987).

2. THERE IS POOR CORRELATION BETWEEN THE RESULTS OF LABORATORY ANIMAL CARCINOGENICITY TESTING IN RATS AND MICE

Freedman and Zeisel select several data sets from the literature and attempt to show that concordance is poor between carcinogenicity results in mice and rats. However, their analysis is misleading in several ways. (1) They regard studies for which only one species was used as indicating disagreement between species. How is it possible for mice and rats to both show carcinogenic effects if only one of the two species is evaluated? (2) They ignore concordance in negative studies. However, a chemical that is not carcinogenic in both rats and mice does indeed show agreement, and such studies must be factored into the overall evaluation. An assessment of association must consider both sensitivity and specificity, i.e., prediction of both carcinogenicity and noncarcinogenicity. (3) They selectively utilize data which they regard as supporting their point of view, while basically ignoring other more comprehensive studies.

One of the studies they chose not to discuss is Purchase (1980), who evaluated the carcinogenicity results for rats and mice for 249 chemicals, finding the following association:

		Rats	
		+	−
Mice	+	111	21
	−	17	100

This association ($p < 0.0001$) is quite impressive. Among the 249 chemicals, 211 (85%) produced a similar response. Sensitivity and specificity ranged from 83–87%.

Rather than citing this more definitive study, Freedman and Zeisel instead attempt to carry out their own evaluation of two subsets of the carcinogenicity studies carried out by the National Cancer Institute (NCI) and the National Toxicology Program (NTP). Their resulting evaluation and interpretation of these data are misleading for the reasons indicated above, but more recent data are available from these studies in any case. A paper by Haseman and Huff (1987) evaluates the association in carcinogenic response in rats and mice for all NCI/NTP studies carried out to date, which include as subsets both of the data sets considered by Freedman and Zeisel. The following summarizes these results:

		Rats	
		+	−
Mice	+	67	36
	−	32	131

These results are somewhat similar to those reported by Purchase, although the overall concordance is lower (198/266, 74%); sensitivity and specificity ranged from 68–78%.

From these two large data sets, which represent a sizable proportion of all laboratory animal carcinogenicity studies carried out to date, it is clear that rats and mice show a strong association in carcinogenic response. It is ironic that other authors, evaluating basically the same studies considered by Freedman and Zeisel, reached the opposite conclusion, namely that the agreement in carcinogenicity results between rats and mice was so complete, that the mouse was "redundant" and need not be utilized in laboratory animal studies (Schach von Wittenau and Estes, 1983)! Although this point of view is also somewhat extreme, clearly the interspecies correlation in carcinogenic response between rats and mice is much higher than that suggested by Freedman and Zeisel. Haseman and Huff (1987) discuss this matter in greater detail.

3. THE RESULTS OF LABORATORY ANIMAL CARCINOGENICITY STUDIES DO NOT CORRELATE WELL WITH HUMAN DATA

Freedman and Zeisel then attempt to demonstrate that laboratory animal and human carcinogenicity results show little correlation. In their evaluation, they utilize data from the International Agency for Research on Cancer (IARC), but present a misleading analysis.

Perhaps the most serious deficiency of their evaluation is their decision to regard the IARC categorization of "limited evidence" of carcinogenicity as equivalent to "no evidence" or "inadequate evidence" of carcinogenicity. This is not appropriate, because chemicals falling into the limited evidence category have shown evidence of carcinogenic effects. For laboratory animal studies, limited evidence often indicates that positive effects have been found, but only in one species. For humans, limited evidence implies that a positive association has been observed between exposure to the agent and cancer, but chance, bias or confounding could not be ruled out with reasonable confidence.

These outcomes are quite different from those associated with inadequately tested chemicals or negative studies. Thus, it is more than merely "unconventional" (Freedman and Zeisel's term), it is INAPPROPRIATE to combine "limited," "inadequate" and "no evidence" categories for purposes of statistical analysis.

Table 7 is perhaps the most important table in the Freedman and Zeisel paper. However, it is misleading, because it pools the limited and inadequate categories. A complete breakdown of the Table 7 data gives these results:

Level of Evidence (Humans)	Level of Evidence (Animals)		
	Sufficient	Limited	Inadequate
Sufficient	13	6	2
Limited	7	2	2
Inadequate	41	33	27

An examination of this table reveals several interesting results that Freedman and Zeisel do not report. For the majority (76%, 101/133) of chemicals, there is simply inadequate human data for a meaningful evaluation of carcinogenicity to be made. Despite Freedman and Zeisel's implication that such studies suggest lack of carcinogenicity, a more appropriate approach is to delete these chemicals from consideration.

Among the remaining 32 chemicals which show evidence (limited or sufficient) of carcinogenic effects in humans, all but four show supporting data in laboratory animals as well. These four "exceptions" are arsenic and certain arsenic compounds (which have since been shown to provide limited evidence of carcinogenicity in animals; see Wilbourn et al., 1986) and three chemicals that have been inadequately tested for carcinogenicity in laboratory animals: conjugated estrogens, chloramphenicol and dienestrol.

Thus, greater than 90% (29/32) of the known or suspected IARC human carcinogens considered by Freedman and Zeisel (and 100% of those adequately tested in laboratory animals) have shown carcinogenic effects in laboratory animal studies as well. This compares to approximately 50% of the more than 300 chemicals evaluated for carcinogenicity by the NCI/NTP (see Haseman et al., 1987). Although the IARC evaluations consider all available data, which may include experiments in addition to the NCI/NTP studies, these results nevertheless suggest that a disproportionate number of known or suspected human carcinogens are positive in laboratory animal studies when compared to the broader universe of chemicals evaluated for carcinogenicity. Clearly, Freedman and Zeisel's conclusion that the IARC data summarized above provides "decisive evidence" that "carcinogenicity in laboratory animals is poor evidence for an effect in humans" is a misrepresentation of these data.

One should not over-interpret the implications of the high sensitivity noted above. Certainly everyone would agree that the proposition that "all human carcinogens are animal carcinogens" does not logically imply that "all animal carcinogens are human carcinogens." However, as is evident from the IARC data, this may be reflecting the fact that far more information is known regarding the carcinogenic potential of

chemicals in laboratory animals than is the case for humans.

As epidemiological data become available, the carcinogenic potential of laboratory animal carcinogens is frequently confirmed in humans (Tomatis, 1979), often in the same target organ (Wilbourn et al., 1986). It is interesting that many critics of laboratory animal studies as predictors of human carcinogenicity are nevertheless strong advocates of laboratory animal experiments to predict the efficacy of potential therapeutic agents and to study basic mechanisms of physiology, biochemistry, toxicology and carcinogenesis.

The major problem in determining an accurate assessment of the overall association between laboratory animal and human carcinogenicity results is not sensitivity, but specificity. It is difficult to calculate a meaningful estimate of the specificity of laboratory animal studies because there are few if any chemicals that have been studied thoroughly enough to be categorized as definitive human noncarcinogens.

Finally, in their evaluation of the IARC data, Wilbourn et al. (1986) concluded that "for many exposures causally related to human cancer, there is at least a target organ in common between humans and at least one animal species." It is interesting that Freedman and Zeisel chose to cite this particular paper to support the opposite conclusion, namely that "animal experiments do not predict the sites that will be affected in humans." Apparently Freedman and Zeisel would accept only virtually 100% concordance as agreement, because the similarity of target sites in animals and in humans for the data presented by Wilbourn et al. (1986) is impressive, especially when one considers the many potential sites of carcinogenicity as noted by Freedman and Zeisel when expressing concern regarding the "multiple end point problem." As indicated earlier, this is one of several examples where Freedman and Zeisel "reinterpret" data of previously published papers to reach opposite conclusions to those presented by the original investigators.

Freedman and Zeisel present several specific recommendations for improvements in the design, analysis and interpretation of laboratory animal carcinogenicity studies. It is doubtful that all of these would be accepted by the scientific community. These recommendations, which reflect an over reliance on statistical considerations at the expense of biological concerns, include the following:

> Recommendation I: Randomly assign animals to cages and to treatment groups; this includes random allocation of cages, because "position in the rack seems to be a risk factor for cancer" and such a randomization would eliminate variability associated with "conditions of husbandry."

The principle of proper randomization is basic to any experimental design, and laboratory animal carcinogenicity studies are no exception. From strictly a statistical point of view, it may indeed be optimal to assign cages of dosed and control animals randomly to locations in the animal room, but from a practical point of view, most investigators would object to this approach because of potential errors in dosing and other related problems associated with a completely randomized housing scheme. Perhaps a reasonable compromise is to employ an experimental design similar to that utilized by the NTP, which consists of random allocation of animals to cages, random assignment of columns of cages in a rack to dosed or control groups and periodic rotation of cage location (NTP, 1984).

Despite the practical difficulties associated with a completely randomized housing scheme, such a design would perhaps be warranted if it could be demonstrated that there were important systematic factors associated with the animal room environment that exerted a strong influence on tumor incidence. However, such is not the case. Some investigators have reported that differences in tumor incidence may be related to cage shelf level, but such associations have either been somewhat inconsistent (Lagakos and Mosteller, 1981) or quite subtle, requiring large sample sizes to attain statistical significance (Greenman, Kodell and Sheldon, 1984). Young (1987) recently reported that "local room effects" appeared to influence the incidence of hepatocellular neoplasms in one experimental group of male mice in the NTP eugenol study (NTP, 1983), but Haseman (1988), in a broader evaluation of all NTP studies, found that the overall frequency of such "significant effects" was similar to chance expectation. Finally, Haseman, Winbush and O'Donnell (1986) evaluated 18 long term carcinogenicity studies in male and female mice and rats that utilized two concurrent control groups housed in separate locations in the animal room, and found that the number of significant differences in tumor incidence between the two control groups agreed closely with the number expected by chance. These results suggest that position in the rack and/or cage effects are not major contributing factors in the determination of tumor incidence.

> Recommendation II: Base statistical evaluations on the proportion of tumor-bearing animals (all sites combined) rather than considering site-specific effects.

The authors admit that this recommendation is "contrary to standard practice in the field" which calls for site-specific analyses. This is somewhat of an understatement because virtually all major guidelines that have been formulated for laboratory animal car-

cinogenicity studies (from industry and government alike) recommend site-specific analyses (Haseman et al., 1986). There are several reasons for this. (1) The high spontaneous rate of certain tumors markedly reduces the sensitivity of a tumor analysis based on "all-sites combined" for detecting biologically important site-specific effects, and (2) pooling a hodgepodge of biologically unrelated tumors does not result in a meaningful variable for biological or statistical analysis. Epidemiological studies of human cancer also focus on site-specific effects.

Freedman and Zeisel argue that pooling of all tumors should be done because otherwise the "multiple end point problem is quite serious" and likely to produce statistical false positive results. In their view, this issue is so important that considerations of study sensitivity and the scientific appropriateness of pooling tumors must be set aside. Indeed, they consider the multiple end point problem so pronounced that it "compromises all bioassay results, whether apparently significant or apparently insignificant."

This is a questionable position for several reasons. First, the primary considerations regarding the pooling of tumors should be biological, not statistical. Even if the multiple end point issue were a major problem (and the authors provide no evidence that this is the case other than making the obvious observation that multiple tumor sites are examined), this would not be sufficient justification to recommend a biologically inappropriate procedure merely because it makes the statistical evaluation easier and reduces the likelihood of a false positive result.

Further, despite Freedman and Zeisel's view that ALL bioassay results are compromised by the "multiple end point problem," the false positive issue is well-recognized by investigators in the area, and taken into account in the overall evaluation of carcinogenicity results. Interpretation of these studies involves far more than a mere compilation of multiple p values and selecting the most significant one; biological factors must also be considered. These include the historical control rate of the tumor of interest, whether or not supporting non-neoplastic effects were observed, and whether or not the apparent carcinogenic response was supported by similar effects in other sex-species groups. Thus, the International Agency for Research on Cancer (Gart et al., 1986) cites a number of studies that have demonstrated that "rules which attempt to model the actual decision process indicate that the false-positive rates are close to the nominal level." The Office of Science and Technology Policy (1985) reaches a similar conclusion.

Recommendation III. Much more care is needed in defining end points before the experiment starts. The type of lesion which will be taken as evidence for carcinogenicity should be decided in advance of the study.

The exact meaning of this recommendation is unclear. Because all tumor types are routinely subjected to statistical analysis in laboratory animal studies, what Freedman and Zeisel seem to be suggesting is that this procedure is too broad and should be modified so that consideration is given only to those specific tumor types regarded as relevant end points for the particular chemical under study, and that these tumor types should be specified in advance. However, it is simply not possible in most cases to predict precisely what carcinogenic effects will be observed in laboratory animal studies. It makes no sense to discount an obvious site-specific carcinogenic effect merely because the investigator was unable to specify the target organ and tumor type in advance.

Recommendation IV: Blind "necropsy work" should be employed routinely in laboratory animal studies.

Blind pathology (i.e., histopathological examination without prior knowledge of whether the tissues are from dosed or control animals) has traditionally been a source of debate between pathologists and statisticians, and Freedman and Zeisel reflexively advocate routine use of this procedure to eliminate possible bias. As noted earlier, in the view of Freedman and Zeisel, "pathologists see themselves as professionals exempt from bias." However, few scientists share this uncharitable view of pathologists, and even the American Statistical Association has softened its stance on blind pathology. In a position paper soon to be published in response to an editorial by the American College of Veterinary Pathologists, the American Statistical Association re-emphasizes the importance of blind pathology during the quality assurance and pathology review phases of the study, to eliminate bias and to verify all possible chemically related effects under a more rigorous experimental protocol. However, they agree with the American College of Veterinary Pathologists that routine use of blind pathology at all phases of the histopathology examination, as recommended by Freedman and Zeisel, may not be necessary. The American Statistical Association position is a reasonable compromise between totally open and totally blind pathology.

Recommendation V: Pool carcinogenicity results across sexes.

Not only do Freedman and Zeisel recommend an analysis based on overall tumor incidence, they also recommend pooling results across sexes. Although this is not a major issue, it is unlikely that this particular recommendation will gain widespread acceptance in

the field, particularly since, as noted by Freedman and Zeisel themselves, often "there are marked differences in carcinogenicity across sexes."

Because the authors' initial involvement in the area was apparently studying the possible carcinogenicity of DDT, it is not surprising that a large portion of their paper is devoted to an evaluation of these data. Citing results from a number of laboratory animal studies, Freedman and Zeisel conclude that the evidence for the carcinogenicity of DDT is "quite flimsy" and "taking the results of the bioassays at face value, DDT seems on balance to inhibit tumor development." Thus, they dismiss the 18 experiments in mice and the 3 experiments in rats that produce significant increases in liver tumors on the basis that "this is something of a controversial area."

Freedman and Zeisel quote portions of the Office of Science and Technology Report (1985) that deals with the interpretation of mouse liver tumors. However, they fail to note that the Office of Science and Technology Report concludes that despite certain diagnostic problems "there is consensus that the mouse liver model in principle does have significance in terms of human risk." Thus, Freedman and Zeisel's conclusion that striking liver tumor effects in 21 experiments from 8 different DDT studies provides only "flimsy" evidence of carcinogenicity would certainly seem to be questionable.

Freedman and Zeisel further note that in the DDT studies there were significant decreases as well as increases in tumor incidence. Unfortunately, however, their statistical evaluation of these data is suspect, because they were unable to adjust properly in most instances for survival differences among groups. This is particularly important because at least half of the "significant decreases" in tumor incidence were associated with markedly reduced survival and were likely not "real" effects. It is surprising, in light of the authors' insistence on statistical rigor in other aspects of laboratory animal carcinogenicity studies, that they themselves would carry out statistical analyses of tumor incidence in DDT studies that provided insufficient information to allow proper adjustment for survival differences.

The authors also present a misleading interpretation of the variability of the risk estimates obtained by Haseman and Hoel (1979) for DDT using a multistage model. Freedman and Zeisel report that for liver tumors these authors found that the variation in risk estimates among 15 DDT studies was as high as a factor of 250. Although this is technically correct, what Freedman and Zeisel fail to disclose is that for 13 of the 15 studies, the estimates were within a factor of 10, and for 14 of the 15 were within a factor of 44. That is, only 1 of the 15 studies produced risk estimates that differed greatly from the others. Thus, the overall agreement was quite good, contrary to the impression given by Freedman and Zeisel.

Freedman and Zeisel conclude by citing quotations from various scientific organizations (e.g., the National Academy of Sciences; the International Agency for Research on Cancer; The President's Office of Science and Technology Policy; The Office of Technology Assessment of the United States Congress) which they regard as presenting the "establishment" point of view. Although many statements are taken out of context and the implications misinterpreted, this citation of comments from prominent scientific organizations does have certain useful features: (1) it indicates that these groups do have an awareness of the limitations of qualitative and quantitative risk assessment, and further that (2) these groups conclude that despite these limitations, in the absence of extensive epidemiology data, careful evaluations of laboratory animal studies remain the most definitive means of addressing the issue of human health risk. None of these organizations appear to share the extreme views of Freedman and Zeisel on these issues.

My own view of the value of laboratory animal carcinogenicity studies is perhaps best summarized by the International Agency for Research on Cancer, an agency that Freedman and Zeisel acknowledge is "well respected and draws working groups of scientists from all over the world." In their revised Preamble to its Monograph Series, the IARC (1987) notes that all 37 known or suspected human carcinogens that have been adequately tested in laboratory animals have also produced cancer in at least one animal species. The IARC concludes that "Although this association cannot establish that all agents that cause cancer in experimental animals also cause cancer in humans, nevertheless, in the absence of adequate data on humans, it is biologically plausible and prudent to regard agents for which there is sufficient evidence of carcinogenicity in experimental animals as if they presented a carcinogenic risk to humans." I suspect that this also represents the position of the majority of scientists working in the area.

Although the focus of these comments is primarily on qualitative risk assessment, one comment on quantification seems appropriate. Admittedly, there are uncertainties associated with the use of mathematical modeling in regulatory decision making. However, contrary to the view of Freedman and Zeisel, these uncertainties are well recognized and whenever possible are factored into the overall decision making process, which utilizes a weight of evidence approach based on all available scientific data (e.g., data from studies of mechanisms; short term test results; results from epidemiological studies). Rather than utilizing

a scientific weight of evidence approach, Freedman and Zeisel propose an alternative solution: "informal argument." However, when scientific evaluations of possible human risk are replaced by a lawyer's debating skills, and regulatory decisions determined solely by who is able to "informally argue" their position more persuasively, then the ultimate loser may be our nation's public health.

ADDITIONAL REFERENCES

FREIREICH, E. J., ET AL. (1966). Quantitative comparison of toxicity of anticancer agents in mouse, rat, hamster, dog, monkey, and man. *Cancer Chemother. Reports* **50** 219–244.

GART, J. J., ET AL. (1986). *The Design and Analysis of Long-term Animal Experiments. IARC Sci. Publ.* **79**. IARC, Lyon.

GREENMAN, D. L., KODELL, R. L. and SHELDON, W. B. (1984). Association between cage shelf level and spontaneous and induced neoplasms in mice. *J. Nat. Cancer Inst.* **73** 107–113.

HASEMAN, J. K. (1988). Lack of cage effects on liver tumor incidence in B6C3F1 mice. *Fundam. Appl. Toxicol.* **10** 179–187.

HASEMAN, J. K., ET AL. (1987). Comparative results of 327 chemical carcinogenicity studies. *Environ. Health Perspect.* **74** 229–235.

HASEMAN, J. K. and HUFF, J. E. (1987). Species correlation in long-term carcinogenicity studies. *Cancer Lett.* **37** 125–132.

HASEMAN, J. K., WINBUSH, J. S. and O'DONNELL, M. W., JR. (1986). Use of dual control groups to estimate false positive rates in laboratory animal carcinogenicity studies. *Fundam. Appl. Toxicol.* **7** 573–584.

HUFF, J. E. and MOORE, J. A. (1984). Carcinogenesis studies design and experimental data interpretation/evaluation at the National Toxicology Program. In *Industrial Hazards of Plastics and Synthetic Elastomers* 43–64. Liss, New York.

IARC (1987). Preamble. *On the Evaluation of the Carcinogenic Risk of Chemicals to Humans. Monograph Supplement* 6. Lyon. To appear.

NATIONAL TOXICOLOGY PROGRAM (1983). Carcinogenesis studies of eugenol in F344/N rats and B6C3F1 mice (feed studies). NTP Technical Report 223, United States Dept. Health and Human Services, National Toxicology Program. Research Triangle Park, N. C.

NATIONAL TOXICOLOGY PROGRAM (1984). General statement of work for the conduct of acute, fourteen-day repeated dose, 90-day subchronic, and 2-year chronic studies in laboratory animals. National Toxicology Program, Research Triangle Park, N. C.

RALL, D. P. (1979). The role of laboratory animal studies in estimating carcinogenic risks for man. In *Carcinogenic Risks—Strategies for Intervention* (W. Davis and C. Rosenfeld, eds.). *IARC Sci. Publ.* **25**. IARC, Lyon.

RALL, D. P., ET AL. (1987). Alternatives to using human experience in assessing health risk. *Ann. Rev. Public Health* **8** 355–385.

SCHACH VON WITTENAU, M. and ESTES, P. C. (1983). The redundancy of mouse carcinogenicity bioassays. *Fundam. Appl. Toxicol.* **3** 631–639.

YOUNG, S. S. (1987). Are there local room effects on hepatic tumors in male mice? An examination of the NTP eugenol study. *Fundam. Appl. Toxicol.* **8** 1–4.

Comment

Suresh H. Moolgavkar and Anup Dewanji

"The Emperor has no clothes," say Freedman and Zeisel, and they marshal an impressive array of facts to support their conclusion.

That carcinogenic risk assessment is of enormous practical importance is a truism. It is equally obvious that, if possible, risk assessment should be performed within the framework of a biologically based model of the carcinogenic process. Indeed, if risk is to be extrapolated from high doses to low or from one species to another, then it is absolutely imperative that a biologically based model be used if the results are to be meaningful. Among the models considered by Freedman and Zeisel, the multistage model, first proposed by Armitage and Doll, is the only one that could, even remotely, be considered to be biologically based. Before discussing this model, we would like to state our view of the role of carcinogenesis modeling in the process of risk assessment. A biologically based cancer model is only one component of a rational risk assessment strategy. Such a cancer model relates fundamental biological processes at the cellular level to the incidence of tumors of specific tissues in human or animal populations. In order to relate the parameters of the model to the specific agent under investigation it is necessary to know the dose of the active metabolite responsible for the carcinogenic action in the tissue of interest. Thus, before the carcinogenesis model can be used, either direct measurements of the metabolite in question must be made in the tissue of interest or a pharmacokinetic model must be used to

Suresh H. Moolgavkar is Professor of Epidemiology and Biostatistics, University of Washington, and Member, Fred Hutchinson Cancer Research Center, 1124 Columbia Street, Seattle, Washington 98104. Anup Dewanji is Staff Scientist, Fred Hutchinson Cancer Research Center.