## ADDITIONAL REFERENCES

ARMITAGE, P. and DOLL, R. (1954). The age distribution of cancer and a multistage theory of carcinogenesis. *British J. Cancer* **8** 1–12.

BROWN, K. G. and HOEL, D. G. (1986). Statistical modeling of animal bioassay data with variable dosing regimens: Example—vinyl chloride. *Risk Anal.* **6** 155–166.

DEDRICK, R. L. (1985). Application of model systems in pharmacokinetics in risk quantitation and regulatory policy. Banbury Report No. 19, 187–198. Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.

KNUDSON, A. G. (1985). Hereditary cancer, oncogenes, and antioncogenes. *Cancer Res.* **45** 1435–1443.

KREWSKI, D., MURDOCH, D. J. and DEWANJI, A. (1986). Statistical modeling and extrapolation of carcinogenesis data. In *Modern Statistical Methods in Chronic Disease Epidemiology* (S. H. Moolgavkar and R. L. Prentice, eds.). Wiley, New York.

MOOLGAVKAR, S. H. (1978). The multistage theory of carcinogenesis and the age distribution of cancer in man. *J. Nat. Cancer Inst.* **61** 49–52.

MURDOCH, D. J., KREWSKI, D. and CRUMP, K. S. (1987). Mathematical models of carcinogenesis. In *Cancer Modelling* (J. R. Thompson and B. W. Brown, eds.). Dekker, New York.

THORSLUND, T. W., BROWN, C. C. and CHARNLEY, G. (1987). Biologically motivated cancer risk models. *Risk Anal.* **7** 109–119.

WILSON, J. D. (1986). Time for a change: Guest editorial. *Risk Anal.* **6** 111–112.


# Comment: The Use of Animal Experiments in Cancer Risk Assessment

**J. Kaldor and L. Tomatis**

Among the agents which have been demonstrated to cause cancer in humans, environmental chemicals have been the major focus of public and regulatory concern over the past several decades. In the case of artificially synthesized compounds, particularly strong sentiments can be aroused: why should proven, or even suspected causes of cancer, be allowed to contaminate the human environment? The answer often given is that many manmade carcinogens have become of substantial economic benefit, and the risks which they represent are to be weighed against the costs of their removal from the environment. It has long been recognized that the quantification of both risk and benefit presents major difficulties, and Freedman and Zeisel are only the most recent commentators to conclude that, in particular, the quantitative assessment of cancer risk entails a number of biological assumptions that have not been verified empirically. In reaching this conclusion, with which we are in general agreement, Freedman and Zeisel sidestep into areas where they appear to be out of their depth, and manage to distort several important issues concerning the use of animal experiments as indicators of potential human hazard.

The sine qua non for establishing a chemical as a human carcinogen is undoubtedly the epidemiological finding that, after eliminating as far as possible the effects of bias, confounding or chance, those exposed to the chemical have a higher risk of cancer than those unexposed. Although laboratory experiments and information on chemical structure, metabolism and pharmacokinetics all play an important role in the evaluation process, some grounds can always be found to challenge their relevance if epidemiological data are absent.

The current reality, however, is that for most environmental chemicals, studies of cancer risk in exposed humans are not available. In a few cases, this can be attributed to a lack of resources for studying an appropriate, exposed population. For the overwhelming number of chemicals, however, such a population does not exist, either because it is too difficult to define who is exposed, because when exposure does occur it is in conjunction with other chemicals or because not enough time has elapsed to be able to evaluate the cancer risk in exposed populations. For a number of major classes of chemicals such as the $N$-nitrosamines, which are both animal carcinogens and widespread in the human environment, direct epidemiological evidence of carcinogenicity is still lacking.

In this situation, the process of cancer risk assessment, whether qualitative or quantitative, must rely entirely on experimental data, perhaps combined with some information on human metabolism or other relevant parameters. Although each chemical should be evaluated independently, it is only on examination of results across chemicals that one can evaluate the validity and long term performance of experimental data in predicting human risk. In their critique of risk

*J. Kaldor is a Biostatistician at the Unit of Biostatistics Research and Informatics, and L. Tomatis is Director, International Agency for Research on Cancer, 150 Cours Albert Thomas, 69372 Lyon Cedex 08, France.*

assessment, Freedman and Zeisel question the quality of animal carcinogenicity bioassays, the qualitative relationship between carcinogenicity in animals and in man, and finally the possibility of using bioassay data to quantitatively predict human cancer risk. Each of these issues is discussed in turn below.

## ANIMAL CARCINOGENESIS EXPERIMENTS

The carcinogenicity bioassay evolved from the carcinogenesis experiments that were originally carried out as a part of basic research into the causes of cancer. The evolution into routine safety testing has generally brought about substantial improvements in their design, analysis and reporting, culminating in the high scientific standards of the U. S. National Toxicology Program (NTP) series of assays (Haseman et al., 1987). Even this program has had its difficulties over time and has occasionally suffered major lapses (Smith, 1979). Overall, the standard of data from animal cancer tests is very uneven. Nevertheless, for many chemicals whose carcinogenicity is in question, one must accept less than perfect experiments as the only source of data. Their potential limitations have long been recognized, and indeed are carefully assessed when experimental data are used to evaluate a chemical's carcinogenicity in contexts such as the IARC monographs.

The available experiments on DDT and its main metabolites are no exception. Although some of the tests to which these compounds were submitted were of a standard equal to any carried out at that time, others were certainly deficient in various respects. Freedman and Zeisel chose to ignore these differences in quality in their presentations of all the experimental data analyzed "in a uniform way." Even if it were reasonable to ignore these differences, the substantial variation in the degree to which the test compounds shorten the animals' life-span should have been an immediate contraindication for any statistical method that does not take into account survival time, such as the simple Armitage trend test for proportions applied by Freedman and Zeisel.

Recent observations on the mechanisms underlying liver tumor induction in mice seem to strengthen their significance as predictors of human risk (Reynolds et al., 1987).

## QUALITATIVE CORRELATION OF CARCINOGENICITY

Animals and humans are certainly different in many fundamental ways, but so far, no clear evidence has emerged that a carcinogen in one species does not cause cancer in other species. In the IARC monographs program, as of Volume 41 (1986), the epidemiological data for 44 chemicals, mixtures or groups of chemicals have been evaluated as either "sufficient" or "limited" evidence for cancer causation (Wilbourn et al., 1986). Of this total, 7 had not been tested adequately in animal studies, and of the remaining 37, the animal results provided either sufficient (27) or limited (10) evidence of carcinogenicity. In most cases, the main reason for the limited evaluation in animals was the lack of multiple experiments, not the consistent finding of a negative result. Considering only the chemicals for which there was sufficient evidence of carcinogenicity in humans produces virtually identical results. Thus, discarding chemicals for which there is insufficient data, the sensitivity of animal experiments in detecting human carcinogens approaches 100%. This figure contrasts sharply with the 59% arrived at by Freedman and Zeisel, using rather selective criteria (Table 6 of their paper).

It is true that experimentalists have gone to great lengths to prove carcinogenicity for substances that have been demonstrated to cause cancer in humans, and that the sensitivity of animal bioassays is perhaps overestimated. Nevertheless, the fact remains that no exception to the rule has been found.

Freedman and Zeisel make a much more fundamental, and one is tempted to say intentional, error in their Table 7. The vast majority of the 79% of sufficient evidence animal carcinogens for which human data are "insufficient" to demonstrate carcinogenicity have not even been evaluated in humans, for the reasons mentioned above. It is completely inappropriate to equate an absence of data with a clear negative finding. This error appears in other places in the paper, including the negative classification of experiments quoted in Table 5.

The observation that some chemicals tested by the NTP are only positive in one species is fully consistent with the possibility that species differ in response, but could also be simply a reflection of low power, as Freedman and Zeisel seem to realize. On the other hand, they seem to have missed the point of the argument of Bernstein et al. (1985), which only concerns quantitative correlation and has no bearing on the validity of the qualitative correlation.

## QUANTITATIVE EXTRAPOLATION OF CANCER RISK

The quantitative extrapolation of risks from animals to humans can only be done if one is prepared to accept a mathematical model for the relationship between exposure and cancer incidence. Although there are many possible models, with very little data to distinguish among them, there is one overriding principle that has guided the choice of model, and that is conservatism. For the few human carcinogenic exposures on which good dose-response data are available, it is not possible to exclude a linear term at low

doses, and indeed, a linear term that is barely perceptible in the observable dose range could still dominate at low doses (Kaldor and Day, 1985). For this reason, the EPA and other bodies have based risk assessment on models that assume linearity at low doses, while fully admitting their pragmatic rather than scientific basis. Scientists are all too aware of the complexities involved to take any model of carcinogenesis literally, but generally accept the need for some sort of standardized quantification of the results from animal experiments (Peto et al., 1984). The principle of conservatism also lies behind the choice of the most sensitive animal species or cancer site to indicate human risk.

Cross-species extrapolation may be facilitated in the near future by new developments in biological dosimetry. By using monoclonal antibodies, biochemists can now detect the reaction products of DNA-damaging agents and cellular DNA, and thus potentially have a much closer measurement of the dose received by a target organ (Berlin et al., 1984; Bartsch, Hemminki and O'Neill, 1988). Studies comparing these measurements on different species should shed light on the dose scale that is appropriate across species.

## CONCLUSION

Freedman and Zeisel make many points that are true, if not novel, but they seem to assume that other scientists have never considered the problems they address. Risk assessment is a complex process that can probably never be automated, but it plays an essential role in an advanced industrial society.

## ADDITIONAL REFERENCES


BARTSCH, H., HEMMINKI, K. and O'NEILL, I. K. (1988). Methods for detecting DNA damaging agents in humans. In *Applications in Cancer Epidemiology and Prevention. IARC Sci. Publ.* **89**. IARC, Lyon.

BERLIN, A., ET AL. (eds.) (1984). *Monitoring Human Exposure to Carcinogenic and Mutagenic Agents. IARC Sci. Publ.* **59**. IARC, Lyon.

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

IARC (1986). *On the Evaluation of the Carcinogenic Risk of Chemicals to Humans: Some Halogenated Hydrocarbons and Pesticide Exposures. Monograph* **41**. Lyon.

KALDOR, J. M. and DAY, N. E. (1985). The use of epidemiological data for the assessment of human cancer risk. In *Risk Quantification and Regulatory Policy*, Banbury Report No. 19 (D. G. Hoel, F. P. Perera and R. A. Merrill, eds.) 79–87. Cold Spring Harbor Press, Cold Spring Harbor, N. Y.

PETO, R., ET AL. (1984). The $TD_{50}$: A proposed general convention for the numerical description of the carcinogenic potency of chemicals in chronic exposure animal experiments. *Environ. Health Perspect.* **58** 1–8.

REYNOLDS, S. H., ET AL. (1987). Activated oncogenes in B6C3F1 mouse liver tumors: Implications for risk assessment. *Science* **237** 1309–1316.

SMITH, R. J. (1979). NCI bioassays yield a trail of blunders. *Science* **204** 1287–1292.


# Comment

**William DuMouchel**

## 1. INTRODUCTION: WHY I EXPECTED TO DISLIKE THIS PAPER

As set out and exemplified more fully in DuMouchel and Harris (1983), I believe that analysis of animal studies can be used to form and improve numerical estimates of cancer risk to humans. Professors Freedman and Zeisel, in their abstract, claim that this is "well beyond the scope of the scientifically possible." This paper, similar in spirit to Freedman and Navidi (1986), seems to deny that statistical modeling can really help much when up against the horrible complexity of real-world problems.

Their confrontational style is designed to provoke reactions of outrage among statistician true believers. Among many examples, overstatements like those in the Introduction "[at first] we felt—along with every other educated person—that DDT caused cancer," and "routine bioassays have little to do with basic research," will probably be pounced upon by other discussants. Not all species extrapolation methods rely solely or naively on Abbott's formula, as the authors seem to imply in Section 2. My personal favorite occurs at the end of the paper "we ... find informal argument more appealing [than statistical modeling] because it brings uncertainties into the

*William DuMouchel is Chief Statistician at BBN Software Products Corporation, 10 Fawcett Street, Cambridge, Massachusetts 02238.*