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.

open." And all these years I've been telling my students and clients that statistics is the science of dealing explicitly with uncertainty!

The Bayesian statistician must be especially explicit and open about uncertainty in the modeling process because of the need to specify prior distributions. The review of the literature that the authors present in Sections 5 and 6 seems to reflect confusion about the most confusing aspect of classical statistics: hypothesis testing. Tables 5, 6 and 7 all tabulate studies of many chemicals by whether they found sufficient (significant?) evidence of carcinogenicity in animals and/or humans. If one study is positive and another is negative, do they disagree? Not necessarily, of course, because testing is not estimation, and don't forget the double negative (we don't accept the null hypothesis, we just fail to reject it) and power calculations that should be, but are usually not, accompanying every interpretation of a significance test. If Table 9 listed slopes and standard errors, rather than merely the ratios of these two statistics, a more interesting analysis, along the lines of that of DuMouchel and Harris (1983), could be performed.

2. IN YOUR HEART YOU KNOW THEY ARE RIGHT

But my quibbles with the authors are more a matter of emphasis and style, rather than content. They rightly point out that extrapolation of cancer risks across species is based on hardly any mechanistic theory, and fitting a "black box" statistical relationship can only be justified by many observational studies, which often just don't exist. It takes a lot of (missing) theory to build a human risk estimate out of one or two animal studies. Yet, out of perceived necessity due to lack of data, many (perhaps most?) studies estimating human risk have done just that.

DuMouchel and Harris (1983) use a Bayesian model and estimates of prior uncertainty to estimate a human dose-response slope for the risk of lung cancer from exposure to diesel emissions. If the estimation is based on just two animal studies and one human study of a similar chemical, its interval of posterior uncertainty spans six orders of magnitude! Yet, when a set of 37 studies, covering five different biological systems and ten somewhat similar chemicals, is integrated into an analysis using the same basic prior distributions, the interval of uncertainty for the same parameter spans just two orders of magnitude, a useful result. As stated on page 314 of that paper, "Human cancer risk assessment requires data on many agents in many species. In the absence of strong prior information on cancer mechanisms, one good rat study is just not enough." And on page 306: "it is often infeasible to conduct a precise epidemiological study of the human cancer risks of a particular agent. In such cases, the Bayesian model can help to decide which of many other unperformed experiments might be most informative about the agent's carcinogenicity." We need to think more in terms of which ensembles of human and animal studies need to be performed.

These results accord with the sentiments expressed in the conclusion of the present paper: "A feasible—but expensive—program would call for a series of good [but not routine?] bioassays on a representative sample of agents and species, with the object of measuring interspecies differences in sensitivity." In fact, the Bayesian analysis referenced above was based largely on just such a set of bioassays, commissioned for that purpose by the EPA.

Freedman and Zeisel also rightly point out that extrapolations of cancer risk from high-dose studies to the lower-dose exposures typical of human populations depend almost entirely on which model is assumed to hold. This makes it imperative to do the high-dose animal studies and in vitro studies on agents for which reasonably good human data exist, so that a data base of such comparisons can be assembled and used in calculating uncertainties in other human risk estimates.

Finally, I agree with the authors' caveats about the often post hoc selection of which tumor site to use as the basis for interspecies extrapolation. Not that a Bayesian should be uncomfortable with post hoc selection per se. But even the Bayesian can only justify it after prior expectations of relationships among risks at potential tumor sites have been examined and incorporated into a prior distribution, and I don't know of anyone who has attempted that yet.

All in all, my disagreements with the authors seem to be centered on whether the glass containing possibilities for solving the mouse to man problem is half full or half empty (although they may reject my optimistic interpretation of their frankly pessimistic paper). In any case, Professors Freedman and Zeisel should be thanked for their concise review of this problem, and for providing an opportunity for a stimulating discussion.

ADDITIONAL REFERENCE

FREEDMAN, D. and NAVIDI, W. (1986). Regression models for adjusting the 1980 census (with discussion). *Statist. Sci.* **1** 3–39.