Environmental Health Perspectives
Vol. 47, pp. 333-337, 1983

Extrapolation of Animal Carcinogenicity Data: Limitations and Pitfalls

by M. D. Hogan*

Laboratory-generated animal bioassay data often serve as the basis for estimating potential human cancer risk. However, there is no single procedure that has been universally accepted as the method of choice for extrapolating experimentally observed results to the low exposure levels that are generally of public health concern. All of the models proposed to date suffer from various limitations. Therefore, the most prudent approach may be to rely primarily on the more conservative procedures such as linear extrapolation until a better understanding of the biological mechanisms underlying the process of carcinogenesis is attained.

In addition to the choice of an extrapolation model, there are a variety of other factors, such as the incorporation of background cancer rates, the potential for synergistic reactions, differential pharmacokinetic effects and differences in exposure regimen, that can have a significant bearing on the extrapolation of animal carcinogenicity data to man.

Introduction

Due to a lack of meaningful epidemiologic data, many estimates of the human health risk posed by environmental exposures to potential carcinogens are based on results derived from laboratory animal experimentation. The primary source of this experimental data is the lifetime cancer screening study or bioassay.

Typically, the investigator who deals with cancer risk assessment is interested in excess risks of the order 10^{-4} to 10^{-6} or even less. Animal bioassays sensitive to such low levels of response would obviously require a study population of enormous and totally impractical size. Therefore, the investigator who wants to assess potential human cancer risk using laboratory-based data must be able to extrapolate his experimentally observed results to exposure levels of public health concern.

Since the investigator often has no empirical knowledge about the behavior of the test chemical's underlying dose-response curve, he is usually forced to rely on some form of mathematical model to characterize the unknown relationship between exposure and tumor response, especially in the low-dose region of interest. To date, no single model has been accepted as the most appropriate for low-dose extrapolation. A wide variety of approaches have been proposed, including models that attempt to establish an upper bound on the unknown cancer risk, tolerance distribution models, models that presumably reflect a general mechanistic process hypothesized to underlie tumor onset and models based on time to tumor occurrence. Each of these approaches is recognized as having its own inherent limitations.

Extrapolation Models

Linear Model

The (no threshold) linear model is the simplest and certainly one of the most commonly employed of the low-dose extrapolation techniques. Under this model it is assumed that the probability of developing cancer is directly proportional to the administered dose at least over some "low-dose" region. Therefore, if $\hat{p}_t$ represents the observed proportion of animals who developed tumors after

*Biometry and Risk Assessment Program, National Institute of Environmental Health Sciences, Post Office Box 12233, Research Triangle Park, NC 27709.

being exposed to the test chemical at dose level d, then the estimated risk corresponding to any dose level d' within the range $[0,d]$ is given by

$$p' = (\hat{p}_t/d)d'$$

If the dose range $[0,d]$ falls within the convex portion of the true dose-response curve, then it is clear that p' will actually be an upper bound on the unknown underlying cancer risk. Furthermore, some investigators (*1*) have advocated the incorporation of an additional protective factor in the estimation process by replacing $\hat{p}_t$ with some upper confidence limit on $\hat{p}_t$ determined from binomial distribution theory.

While this procedure obviously has a great deal of intuitive appeal, it is not immune to criticism. For example, there are undoubtedly many instances in which its use leads to an unduly conservative estimate of the probability of developing a tumor. On the other hand, if the experimental dose range falls outside the convex portion of the underlying dose-response curve, the projected risk obtained through linear extrapolation could significantly underestimate the true likelihood of tumor onset. In the absence of prior knowledge about the test chemical's behavior, the experimental data may not be sufficient to determine whether the linear extrapolation is likely to underestimate or overestimate the actual risk.

Tolerance Distribution Models

The next major class of models employed in low-dose extrapolation is known as the tolerance distribution models. This title reflects the underlying model assumptions (*2,3*) that every individual (animal) in the study population has his/her own level of tolerance to or threshold for the exposure under investigation, that these threshold levels vary among individual members of the study population and that the distribution of these tolerance levels can be described in terms of some cumulative probability distribution function F. Given these assumptions, the probability $P(d)$ of developing a tumor as a result of exposure to a dose d of the chemical of interest is equivalent to the probability of an individual having a threshold level of d or less, i.e.,

$$P(d) = F(\alpha + \beta \log d)$$

If F is assumed to be the cumulative normal distribution, then $P(d)$ is the log-probit model long employed by biologists in median lethal dose estimation. Alternative models, such as the logistic and extreme value models, can be obtained by adopting other distributional forms for F (*3*).

The use of the log-probit model in low-dose extrapolation of carcinogenicity data was pioneered by Mantel and Bryan (*4,5*). They proposed that an upper bound on the dose associated with some preselected "acceptable" level of cancer risk could be estimated by extrapolating from an upper confidence limit on the proportion of experimental animals exhibiting tumors (plotted on a probit versus log dose scale) along a dose-response line with a fixed slope of one. Their choice of a slope of one was based on empirical observations which suggested that this value would typically lead to a conservative estimate of the dose corresponding to the "acceptable" risk level.

While the Mantel-Bryan procedure was initially regarded by many investigators as a reasonable technique for estimating the low-dose cancer risk, subsequent research has enumerated a variety of deficiencies associated with its application that have raised serious doubts (*2,3,6*) about its continued utility. For example, even though the Mantel-Bryan procedure was supposedly designed to generate a conservative estimate of the dose associated with some predetermined "acceptable" level of risk, the log-probit model upon which it is based tends to approach the origin more rapidly than any of the alternative models employed in low-dose extrapolation. Thus, in the low-dose region, the Mantel-Bryan procedure can produce "safe-dose" estimates that are orders of magnitude higher than those generated by competing techniques (*3*). In addition, there is no mechanistic model for carcinogenesis that is reasonably approximated by a log-probit distribution. Finally, the Mantel-Bryan procedure employs Abbott's correction factor (*4*) to take background cancer incidence into account, which has certain implications about the stochastic independence of background and chemically induced tumors that may not hold in general.

"Hit" Models

From a mechanistic point of view the most interesting class of low-dose extrapolation procedures is made up of what are commonly known as "hit" models, a name that refers to a hypothesized process of carcinogenesis upon which they are based. Essentially, this process depicts cancer originating as a malignant cell that has undergone a finite number of somatic mutations or incurred a finite number of hits or receptor interactions with the study chemical. [This mechanistic representation of carcinogenesis has been discussed in some detail (*7*).] Included in this category are the one-hit model,

$$P(d) = 1 - \exp\{-\lambda d\}$$

where λ is an unknown model parameter and λd is the expected number of hits at dose d; the multistage model,

$$P(d) = 1 - \exp\{-(\lambda_0 + \lambda_1 d + \lambda_2 d^2 + \cdots + \lambda_k d^k)\}$$

where the λ_i are unknown, nonnegative model parameters and k corresponds to the number of transitional events or stages in the carcinogenic process; and the Gamma multi-hit model,

$$P(d) = \int_0^d [\Gamma(k)]^{-1} \lambda^k t^{k-1} \exp\{-\lambda t\} dt$$

where λ is an unknown model parameter and $\Gamma(k)$ is the standard gamma function.

The one-hit model, which assumes that the carcinogenic process can be initiated after a single cellular transition has occurred, is certainly the simplest of the various hit models to apply. It often produces estimates that are quite similar to those obtained with the linear model, but it has the advantage of using all of the experimental data simultaneously in the estimation of its model parameter. However, since the one-hit model involves only a single unknown parameter, it does not always generate an adequate fit to the experimental observations.

The multistage model in its various formulations has been, perhaps, the most intensely researched and widely employed of the different hit models. Complex computer packages have been developed for fitting this model to experimental data, and many regulatory bodies regard it as one of the standard tools of risk estimation and assessment. Nevertheless, a number of objections have been raised against its unrestricted usage. Since it often behaves like a simple linear model in the low-dose region, it tends to be one of the more conservative of the currently popular extrapolation procedures. Furthermore, the multistage model as it is typically applied assumes that background tumor incidence and the chemically induced carcinogenesis are basically additive components of the same underlying process; and some investigators question the appropriateness of this assumption (*8*).

The last of the models to be considered in this category, the Gamma multi-hit model, can be thought of as an extension of the simple one-hit model (*3*) in which at least k hits are required to initiate tumor development. While this procedure has been recommended by the Food Safety Council (*8*) for use in low-dose cancer risk estimation, a recent, detailed assessment (*9*) of the analyses presented in the Council's report indicated that there are a number of serious practical problems associated with its employment. For instance the Gamma multi-hit model can produce relatively high projections of the background incidence rate even when no tumors are observed among the experimental controls. This, in turn, can lead to unreasonably high estimates of the "safe" dose corresponding to some predetermined "acceptable" level of risk. On the other hand, since the Gamma multi-hit model allows for hyperlinearity, it can also generate unrealistically low safe-level estimates. Then, like the log-probit model, it treats the induced tumor rate as being independent of background. Although this assumption is very difficult to verify empirically, it can have a tremendous impact on the magnitude of the low-dose risk estimate.

Time to Tumor Models

Now, each of the extrapolation models considered above assumes that the basic experimental data will be limited to tumor incidence over time. However, in some large-scale studies there may also be useful information available on time to tumor occurrence or observance. Given such data, it may be possible to fit a mathematical model that depicts a more complex relationship between dose, tumor incidence, and age at which the tumor is detected, and then to use this model to develop estimates of the low-dose risk. Among the various procedures that have been applied to these types of data are the log-normal (*1*), Weibull (*1*), Armitage-Doll Multistage (*10*) and Hartley-Sielken (*11*) models.

Unfortunately, the experimental data obtained from the typical cancer bioassay will often not be of sufficient sensitivity or quantity to allow an investigator to distinguish between the "fits" provided by competing models (*10,12*). (In fact in many instances time to tumor information may not even be collected.) Yet, as with the quantal response models, the selection of a specific time to tumor model can have a very significant impact on the estimate of low-dose risk. On the other hand, the existence of a substantial quantity of bioassay time to tumor information may not ensure satisfactory modeling. For example, preliminary analyses of the massive data set generated by the NCTR ED_{01} study (*13*) of 24,000 mice exposed to the known carcinogen 2-AAF suggest that even the relatively sophisticated Armitage-Doll and Hartley-Sielken time to tumor models do not adequately describe the experimental results (*14*).

Even this brief review of carcinogenesis extrapolation models clearly indicates that there is no general agreement as to the method of choice for low-dose risk estimation. Given the various shortcomings associated with each of the procedures under discussion, it is difficult to decide whether these models are really reflecting some underlying biological mechanism or merely acting as curve-fitting devices. [The recently expressed doubts (*15*)

about the role of somatic mutations in carcinogenesis, a role that underlies the different hit models, contribute to this uncertainty.] Until additional research confirms the biological appropriateness of these various models, the most prudent approach from a public health viewpoint may be to rely heavily on the more conservative, simplistic approaches to low-dose risk estimation like that offered by the linear model.

To obtain an empirical estimate of the variability associated with the selection of this (or any other) extrapolation model, it may be worthwhile to consider the point estimates generated by other extrapolation models as well as the magnitude of the range in these projections. (If this exercise is undertaken, the log-probit model should probably be excluded from consideration because of its potential for underestimating the true risk.)

Other Considerations

There are a number of other issues besides the selection of a mathematical model that need to be considered when extrapolating animal carcinogenicity data. As has already been indicated, deciding whether to treat the background and study chemical exposures as independent or additive phenomena can have a very significant impact on the low-dose risk estimates. It is well known (*16*) that if background can be regarded as additive in a mechanistic sense, then, quite generally, the relationship between dose and response will be essentially linear in the low-dose region. However, recent results obtained by Hoel (*17*) imply that if background is not totally independent of the chemical exposure under study, then low-dose linearity will tend to prevail regardless of the model selected for extrapolation. The significance of this finding is obvious when one considers the likelihood that the assumption of total independence can be justified on biological grounds for any given chemical.

Another issue that can have important bearing on the estimation of low-dose risk is the relationship between the externally administered dose and the concentration of active material reaching the target tissue or cell. While the preceding discussion of the various extrapolation models essentially assumed that the fate of the chemical under study does not depend on dose, this assumption will not always be appropriate. One reason for the possible existence of dose dependency is that many potentially carcinogenic compounds require metabolic activation, which is an enzyme-mediated process. Consequently, the amount of reactive metabolite reaching the target tissue will not necessarily be proportional to the administered dose. The value of incorporating pharmacokinetic considerations into the extrapolation process has been demonstrated by both Gehring et al. (*18,19*) and by Anderson et al. (*20*). Each investigator reanalyzed previously published data on vinyl chloride and found that a significantly improved fit to the experimental data was obtained for either the log-probit or multistage model when the pharmacokinetics of the study chemical were taken into consideration.

Investigators can also be confronted with the problem of deciding what is the most appropriate way to use continuous dosing, lifetime bioassay data to estimate human cancer risk associated with an exposure of much shorter duration. Whittemore (*21*) and Day and Brown (*22*) have investigated the general issue of lifetime versus less than lifetime exposure assuming an underlying multistage model. While both studies provided insight into the relationship among different temporal patterns of exposure, the stage of the carcinogenic process effected by the exposure in question, and the resulting cancer risk, the issue is still far from being fully resolved.

Another issue than can affect low-dose extrapolation and that is also far from being completely resolved is the problem of reconciling the single-agent exposure regimen of the standard animal bioassay with the simultaneous multiple exposure environment of man. Usually, very little is known about the potential synergistic/antagonistic effects of the study compound when given in combination with other agents. Therefore, the issue is often ignored; or, at most, simple additivity is assumed.

Finally, any attempt to extrapolate animal cancer data to man must ultimately deal with the issue of species scale-up. Many investigators feel that this is an even more difficult problem to address than low-dose extrapolation. Species extrapolation will obviously be complicated if there are significant differences between the experimental test species and man with respect to important pharmacokinetic effects such as metabolism and excretion. In addition, temporal, size, population structure, and exposure regimen differences can also complicate the scaling-up process.

In spite of all the limitations and pitfalls associated with the extrapolation of animal carcinogenicity data, however, the process is still a more reasonable procedure for estimating human cancer risk than the traditional safety factor approach which it has supplanted. Extrapolation can at least be used to rank-order priorities in addressing the possible carcinogenic hazards posed by various environmental exposures and to develop a rough estimate of the health impact of such exposures.

REFERENCES

1. Hoel, D. G., Gaylor, D. W., Kirschstein, R. L., Saffiotti, U., and Schneiderman, M. A. Estimation of risks of irreversible delayed toxicity. J. Toxicol. Environ. Health 1: 133-151 (1975).
2. Hogan, M. D., and Hoel, D. G. Extrapolation to man. In: Methods in Toxicology (A. W. Hayes, Ed), Raven Press, New York, pp. 711-731.
3. Krewski, D., and Van Ryzin, J. Dose response models for quantal response toxicity data. In: Current Topics in Probability and Statistics (M. Csorgo, D. Dawson, J. N. K. Rao, and E. Saleh, Eds.), North-Holland, Amsterdam, in press.
4. Mantel, N., and Bryan, W. R. "Safety" testing of carcinogenic agents. J. Natl. Cancer Inst. 27: 455-470 (1961).
5. Mantel, N., Bohidar, N. R., Brown, C. C., Ciminera, J. L., and Tukey, J. W. An improved "Mantel-Bryan" procedure for "safety testing" of carcinogens. Cancer Res. 35: 865-872 (1975).
6. Crump, K. S. Response to open query: theoretical problems in the modified Mantel-Bryan procedure. Biometrics 33: 752-757 (1977).
7. Safe Drinking Water Committee. Drinking water and health. National Academy of Sciences, Washington, DC, 1977.
8. Scientific Committee of the Food Safety Council. Quantitative risk assessment. Food Cosmet. Toxicol. 16: 109-136 (1978).
9. Haseman, J. K., Hoel, D. G., and Jennrich, R. I. Some practical problems arising from the use of the gamma multihit model for risk estimation. J. Toxicol. Environ. Health 8: 379-386 (1981).
10. Hoel, D. G. Animal experimentation and its relevance to man. Environ. Health Perspect. 32: 25-30 (1979).
11. Hartley, H. O., and Sielken, R. L. Estimation of "safe doses" in carcinogenic experiments. Biometrics 33: 1-30 (1977).
12. Whittemore, A., and Altschuler, B. Lung cancer incidence in cigarette smokers: Further analysis of Doll and Hill's data for British physicians. Biometrics 32: 805-816 (1976).
13. Cairns, T. The ED_{01} study: introduction, objectives, and experimental design. J. Environ. Pathol. Toxicol. 3: 1-7 (1980).
14. Brown, K. G., and Hoel, D. G. Modeling time-to-tumor data: analysis of the ED_{01} study. Fundamental Appl. Toxicol., in press.
15. Cairns, J. The origin of human cancers. Nature 289: 353-357 (1981).
16. Crump, K. S., Hoel, D. G., Langley, C. H., and Peto, R. Fundamental carcinogenic processes and their implications for low dose risk assessment. Cancer Res. 36: 2973-2979 (1976).
17. Hoel, D. G. Incorporation of background in dose-response models. Fed. Proc. 39: 73-75 (1980).
18. Gehring, P. J., and Blau, G. E. Mechanisms of carcinogenesis: Dose response. J. Environ. Pathol. Toxicol. 1: 163-179 (1977).
19. Gehring, P. J., Watanabe, P. G., and Park, C. N. Resolution of dose-response toxicity data for chemicals requiring metabolic activation: example—vinyl chloride. Toxicol. Appl. Pharmacol. 44: 581-591 (1978).
20. Anderson, M. W., Hoel, D. G., and Kaplan, N. L. A general scheme for the incorporation of pharmacokinetics in low-dose risk estimation for chemical carcinogenesis: example—vinyl chloride. Toxicol. Appl. Pharmacol. 55: 154-161 (1980).
21. Whittemore, A. S. The age distribution of human cancer for exposures of vary intensity. Am. J. Epidemiol. 106: 418-432 (1977).
22. Day, N. E., and Brown, C. C. Multistage models and primary prevention of cancer. J. Natl. Cancer Inst. 64: 977-989 (1980).