inhalation carcinogenesis

Proceedings of a Biology Division, Oak Ridge National Laboratory, conference held in Gatlinburg, Tennessee, October 8–11, 1969

Sponsored by
National Cancer Institute
and
U. S. Atomic Energy Commission

Editors

M. G. Hanna, Jr.
P. Nettesheim
J. R. Gilbert

April 1970

U. S. ATOMIC ENERGY COMMISSION Division of Technical Information

ACTUARIAL METHODS IN THE EVALUATION OF DATA FROM LONG-TERM ANIMAL EXPERIMENTS

FRANCIS J. C. ROE
Chester Beatty Research Institute, London, England

A distinction has to be made between qualitative and quantitative experiments. An extreme example of the former would be if, after the administration of a new drug, a mouse stood on its head for 5 minutes. It might be wise to repeat such an experiment, but insofar as the knowledge and experience of the investigator permitted him to regard the observation as "unique" and a form of behavior that has not previously been recorded, formal statistical evaluation would be superfluous.

Perhaps a more serious matter is the inappropriate use of statistics in experiments in which animals have not been adequately randomized between treatment and control groups initially. Proper randomization of animals is a prerequisite for a quantitative experiment in which statistical evaluation is contemplated.

Equally important, particularly in the case of life-span studies, is that animals should be observed on every day of the week including Saturday and Sunday. Postmortem autolysis occurs rapidly in mice and, in all species, detailed histopathological evaluation is rendered increasingly difficult as the interval between death and necropsy increases. In long-term mouse studies, a few mice (up to 5%) are usually "lost" for the purpose of pathological evaluation despite the utmost care and 7-day per week observation, but for every day of the week that the laboratory is closed, a further 15% of the animals are likely to be rendered unavailable for pathological study. It is common to see salvage rates of the order of only 66% for mouse experiments conducted in laboratories which are closed on Saturdays and Sundays. Whatever statistical analysis is applied to the results of such experiments, the answer obtained may be misleading.

In our laboratory we kill sick animals, rather than wait for them to die, because we believe that it is more important to be able to make a sound pathological evaluation than to avoid the possibility that, by killing a sick animal that might have recovered, one would bias the results of an experiment. With

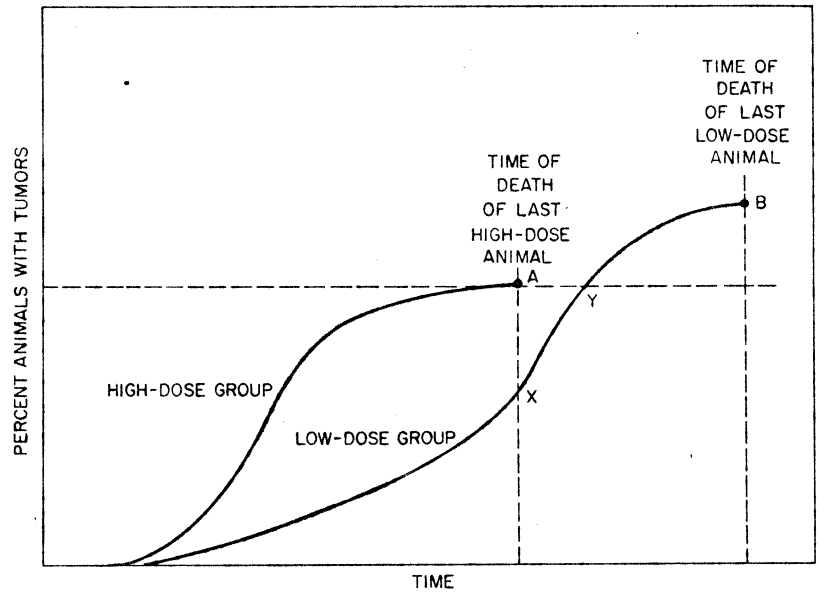

Fig. 1 – Commonly used ways of expressing data from carcinogenicity experiments.

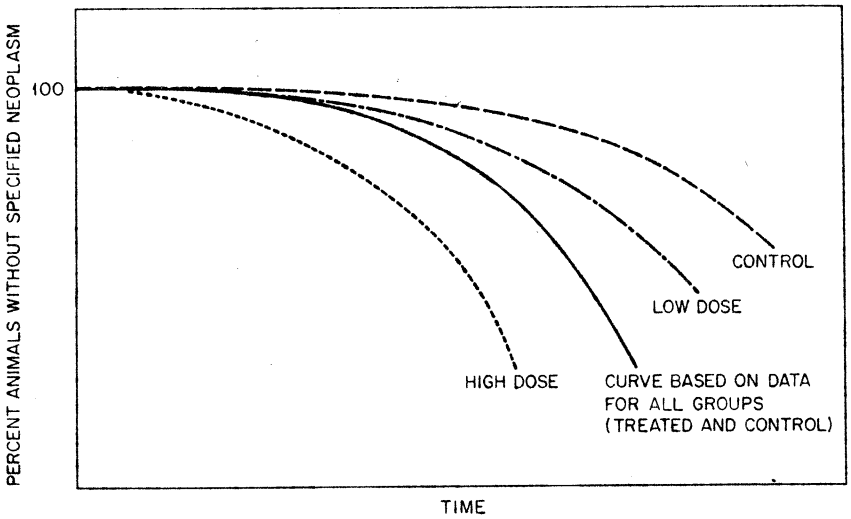

Fig. 2 – Calculation of risk of development of specified neoplasm in differently treated groups. The overall experience of each group can be described quantitatively and, if required, significance levels can be calculated in respect of apparent differences.

deliberately killed animals, we are able to follow a regular necropsy regime. This would not necessarily be possible if animals were permitted to die.

At this conference we have heard speakers refer to "tumor incidence" in response to different forms of treatment, and frequently such reference has been made without mention of comparative survival times. Commonly used ways of expressing data from carcinogenicity experiments are illustrated in Fig. 1. If a test material is toxic it is possible for the final percentage of a group that developed neoplasms to be higher in a low-dose group (B) than in a high-dose group (A). Without qualification with regard to survival, this result would be misleading to say the least. Comparison of percentages of tumor-bearing animals in the two groups at a stated time (*e.g.*, A and X), or of times by which a given percentage of animals in a group have become tumor bearers (A and Y) could also be misleading, and I make a plea for the use of actuarial methods in the evaluation of data from laboratory animal experiments.[1-4] A continuous adjustment for intercurrent (or nonrelevant, *i.e.*, nontumor) deaths is made, and a tumor-free survival curve is constructed which takes the form illustrated in Fig. 2. This curve shows the estimated survival of the animals if tumor development were the only cause of death. The expected shape of a curve for a group of animals depends only on the tumor incidence rates at various ages and not at all on the death rates from other causes, such as toxicity on local epidemic infections. The curves for the individual groups can be compared numerically by the method of Gehan[5]. Each curve is built up on the basis of numbers of animals developing tumors as proportions of animals alive and at risk of doing so at serial points in time. In other words, each curve is an expression of "time-standardized risk." The burden of my argument is simply that we should abandon the concept of percent of animals that develop neoplasms in favor of the concept of "time-standardized risk."

By the use of this method of analysis we are currently obtaining far more information from long-term studies than was previously possible. Sometimes actuarial analysis shows that a conclusion based on crude analysis was completely wrong.

REFERENCES

1. J. O. Irwin and N. Goodman, Statistical treatment of measurements of carcinogenic properties of tar and mineral oils. *J. Hyg.*, 44(1946) 362.
2. C. C. Spicer, Estimation of tumour susceptibility in pure lines. *Brit. J. Cancer*, 1(1947) 298.
3. H. I. Pilgrim and J. E. Dowd, Correcting for extraneous death in the evaluation of morbidity or mortality from tumour. *Cancer Res.*, 23(1963) 45.
4. M. C. Pike and F. J. C. Roe, An actuarial method of analysis of an experiment in two-stage carcinogenesis. *Brit. J. Cancer*, 17(1963) 605.
5. N. Mantel, Ranking procedures for arbitrarily restricted observation. *Biometrics*, 23(1967) 65.