Fd Chem. Toxic. Vol. 31, No. 3, pp. 225–231, 1993

WHAT DOES CARCINOGENICITY MEAN AND HOW SHOULD WE TEST FOR IT?*

Introduction

The main reasons why these questions are important stem from the aim to reduce the existing burden of cancer among humans and to make sure that this burden is not increased when new chemical agents are introduced into the human environment.

Definition of cancer

It follows from this objective that the first need is to define the term 'cancer'. A simple definition of cancer is that it is a disease characterized by the proliferation of abnormal body cells and by the spread of these cells to other tissues by invasion or metastasis. Clusters of proliferating cells that are neither invasive nor capable of metastasizing are not cancers. Thus, benign neoplasms are not cancers. Nevertheless, they are often used as surrogates for cancers by those who conduct carcinogenicity tests in laboratory animals. This is, first, because agents and/or mechanisms that give rise to benign neoplasms also give rise to malignant ones; and secondly, and very regrettably, because pathologists often find it difficult to distinguish between benign and malignant lesions or cannot agree about the distinction. For a time it was being advocated that, insofar as it is not possible to set hard and fast criteria for the distinction between benign adenoma and adenocarcinoma in the rodent lung, or between benign and malignant rodent liver cell tumours, the prudent thing to do was to regard all these lesions as malignant. As should have been anticipated, this solved nothing because pathologists were equally unable to agree on how to distinguish between hyperplastic/metaplastic lesions and neoplasms in the lung, or between foci of hepatocellular change and neoplasia in the liver.

Fortunately, it has once more become respectable for pathologists to diagnose lesions as benign neoplasms, partly as a consequence of pathologists studying individual lesions with a view to agreeing criteria and terminology, and partly because commonsense has, for once, gained the ascendancy.

Surprisingly, perhaps, there has been less argument about the malignancy or non-malignancy of multicentric neoplasms such as those affecting the reticuloendothelial or haematopoietic systems, even though it can be very difficult to distinguish between, say, a slowly progressive form of malignant lymphoma and a generalized lymphocytic proliferative condition of infectious origin.

Invasiveness is a fairly reliable criterion for distinguishing between benign and malignant neoplasma of epithelial cell origin. However, it is important to define what is being invaded. Thus, the extension of a mammary tumour through the mammary fat pad in which it arose is not reliable evidence of malignancy.

When it comes to neoplasms of connective tissue origin, invasiveness can be of limited value as a criterion, because it is a property of many kinds of connective tissue cell that they can migrate through tissues. Thus, a benign dermal fibroma in a rat, which virtually no pathologist would suspect of malignancy, may extend on either side of the panniculus carnosus muscle. By contrast, penetration of this same muscle layer by a tumour arising in the epidermis is regarded as a reliable indicator of malignancy.

Tumour incidence data derived from animal tests versus human cancer mortality data

In the light of the main reasons for conducting laboratory tests for carcinogenicity, it is a disturbing fact that it is often extremely difficult to relate the findings to the assessment of putative cancer risk in man.

The results of animal tests for carcinogenicity are nearly always expressed in terms of comparisons of the percentages of animals in control, low-, medium- and high-dose groups that, when they die or are killed, are found to have one or more benign or malignant tumor(s) of specified kinds. At one time, no attempt was made to correct apparent differences in incidence for between-group differences in survival. Fortunately, survival differences are now normally taken into account. By contrast, the only reliable human data that are available for many forms of cancer are mortality data derived from death certificates. Autopsy rates are generally low (e.g. less than 20%) and many of the autopsies are carried out on young people involved in accidents or on people dying under suspicious circumstances. Autopsy rates among older people dying supposedly from natural causes are much lower than 20%. Furthermore, the autopsies carried out do not necessarily follow a

*This paper is based on a BIBRA seminar given by Dr Roe on 27 October 1992.

systematic procedure and are not necessarily particularly thorough.

For these reasons we have virtually no reliable data for the incidence of small malignant neoplasms or for benign neoplasms, whether small or large, arising internally in humans. Thus, there is simply no way of relating data for such tumours derived from animal studies to cancer risk in man.

To make matters worse, it is clear that human cancer mortality data are themselves exceedingly inaccurate. Comparisons of causes of death based on clinical observations without autopsies and causes of death based on autopsy findings, show major—indeed startling—differences! (Heasman and Lipworth, 1966; Report of Joint Working Party of the Royal College of Pathologists, Physicians and Surgeons, 1991).

Human cancer mortality data versus animal tumour incidence data: can the comparability be improved?

Highly desirable though it may be from a scientific point of view, there is no chance that the present very low human autopsy rates are going to change materially for the better. Nor alas is there any likelihood that, in the near future, the accuracy of clinical diagnosis with regard to the incidence of neoplasia will materially improve even in relation to large fatal internal cancers let alone for small internal malignant neoplasms or for internal benign neoplasms.

One thing that could be attempted, however, is to try to express the results of animal carcinogenicity tests in terms of cancer mortality. Even here there are limits to what can be done. First, whereas death from cancers in humans is often prevented or postponed by treatment, for many obvious reasons this could not be done in animal tests. Secondly, it is not permissible to allow animals literally to die from cancers, or indeed from other debilitating diseases, because this might entail preventable suffering. Thus, it is normal practice to subject sick animals to euthanasia. This being so, 'mortality' needs to include 'debilitation rendering euthanasia necessary'. Thirdly, it is often difficult or impossible to identify the cause of death, or disability requiring euthanasia, in laboratory rats and mice; either one finds no cause or one finds several causes of debilitating illness without being able to identify a single predominant cause. Fourthly, in the case of some endocrine neoplasms, irrespective of whether they are benign or malignant histologically, it may be impossible to determine whether and, if so, to what extent, they contributed to death or the need for euthanasia. This is true, for instance, for neoplasms of the pituitary gland.

For these and other reasons it is not going to be possible to obtain precise cancer mortality data for small laboratory rodents in carcinogenicity tests. On the other hand, it would be possible in many cases to opine whether small neoplasms, irrespective of whether they are benign or malignant, are likely to have contributed to the deaths of animals or for the need for them to be killed by euthanasia. In my opinion this would constitute an improvement on present common practice.

In many laboratories the pathologist is presently required to list 'factors contributory to death'. How he or she does this varies from laboratory to laboratory. Very often a computer selects which factors are listed having been instructed, for instance, to include all neoplasms irrespective of their size or kind. It would be nice to believe that a thoughtful pathologist could do better than this on the basis of careful observations and good judgement!

Yet another problem is the fact that a majority of animals in some studies are simply sacrificed at the end of the 2-yr stint on test. Many of these animals are seemingly in good general health. Nevertheless, in a proportion, neoplasms are found including small malignant neoplasms or the early stages of multifocal lymphoreticular neoplasms. It is not possible to relate these findings to age-standardized cancer mortality.

Overall, then, it would seem that there is a great gulf between the type of cancer data that laboratory experiments can generate and the kind of data that epidemiologists usually use when studying cancer risk to humans. It is not possible to generate reliable cancer mortality data from animal studies and not possible to collect reliable tumour incidence data for internal body sites from human populations.

Definition of carcinogenicity

I define carcinogenicity as 'the enhancement of age-standardized incidence of malignant neoplasia'. Enhancement of benign tumour incidence does **not** constitute carcinogenicity although it may—and often does—provide grounds for 'suspicion of possible or probable carcinogenicity'.

An increased incidence of lesions other than benign neoplasms may also provide grounds for suspicion of possible carcinogenicity. Depending on the tissue/organ, lesions that fall into this category include:

necrosis followed by regenerative hyperplasia [e.g. nasal epithelium—formaldehyde and other aldehydes (Roe and Wood, 1992)].

$\alpha_{2\mu}$-globulin nephropathy in male rats which leads to necrosis and subsequent replacement of proximal renal tubular epithelium [e.g. *d*-limonene, 2,2,4-trimethylpentane (Borghoff *et al.*, 1990 and 1991)].

persistent hyperplasia [e.g. TPA (tetradecanoylphorbol acetate) in mouse skin (Iversen, 1988); pancreatic exocrine cell hyperplasia in response to raw soy protein with trypsin inhibitory activity in the rat (McGuinness *et al.*, 1980)].

cellular atypia and carcinoma *in situ* (e.g. senile keratoses in response to UV radiation in bladder epithelium).

Relevance of mutagenicity

At one time theorists in carcinogenesis came close to suggesting that evidence of mutagenicity constitutes virtual proof of carcinogenic potential. In recent years there has been a widespread withdrawal from this belief, particularly in relation to the interpretation of *in vitro* tests for genotoxicity and in relation to malignant transformation of mammalian cells *in vitro*. On the other hand, positive results in mammalian *in vivo* tests for genotoxicity—either gene mutation or clastogenicity—are still regarded as providing grounds for suspicion of possible or probable carcinogenicity. Most importantly, lack of evidence of genotoxic activity does not establish lack of carcinogenic potential.

Tumor initiation, tumour promotion and the two-stage theory of carcinogenesis

Without question the two-stage theory of carcinogenesis as proposed by Peyton Rous (Rous and Kidd, 1941) and elaborated by Berenblum and Shubik (1947) has been the stimulus for extensive research and has served importantly to advance our knowledge of carcinogenesis. Nevertheless, I would argue that the concept should no longer be regarded as central to understanding in carcinogenesis and that the terms 'tumour initiation' and 'tumour promotion' should now be phased out of use (Roe, 1988). It is clear (e.g. from the mathematical analysis of both epidemiological and experimental data) that more than two stages are usually involved in carcinogenesis (Armitage and Doll, 1961; Peto *et al.*, 1975), and it is clear from observations in animals that a genotoxic stage (i.e. tumour initiation) does not necessarily constitute the first stage (Roe, 1989). It remains highly plausible that at least one genotic event is involved in the genesis of most cancers and likely that multiple (i.e. successive) mutations are sometimes implicated.

Endogenous mutagens

Until recently the thrust of research in the field of cancer prevention has drawn force from the belief that 80% or more of the cancers that arise in people are environmental, as distinct from genetic, in origin.

Comparison of incidences of cancers of different types in different geographical regions and cultures combined with observations on cancer incidence in people who migrate from one geographical area/culture to another are consistent with this belief. On the other hand, the assumptions made by many—first that 80% of cancers could be avoided, and secondly that all 80% are attributable to exposure to xenobiotic carcinogens introduced into the environment by man—are both nonsensical.

Recently, interest has been growing in the fact that DNA-damaging electrophiles are generated in abundance during normal metabolic processes, particularly during the peroxidization of fats (Ames and Saul, 1987). In parallel, attention has been drawn to the fact that during the single-strand phase of the mitotic cycle, DNA repair is impeded with the result that mutation is rendered more likely. This second fact has led to the theory that mitogenesis (increased rate of cell division) predisposes to mutagenesis (increased mutation because of unrepaired DNA damage) (Ames, 1989; Ames and Gold, 1990; Cohen and Ellwein, 1990). Weinstein (1991) has argued against the simplicity of this theory, which rather neatly explains how non-genotoxic irritants that cause persistent hyperplasia or that cause necrosis followed by regenerative hyperplasia, can enhance the risk of mutation and subsequent development of cancer. However, this theory seems to me to be the only one that plausibly explains how simple slight overfeeding dramatically increases the risk of cancer development and how simple slight calorie restriction reduces it (Roe, 1991; Roe and Lee, 1991; Roe *et al.*, 1991).

Non-genotoxic carcinogen

This term has now, at long last, come to be widely accepted as being applicable to agents that enhance the risk of development of malignant neoplasms despite giving negative results in a comprehensive battery of genotoxicity tests. In one sense, the term is misleading because mutation is probably always involved in the sequence of events that leads up to cancer development, but the point is that in the case of a non-genotoxic carcinogen the mutation is not caused by the agent itself nor by any of its metabolites. Acceptance of the term was long resisted by researchers who had spent their lives studying metabolic activating mechanisms or developing ingenious tests for detecting DNA damaging potential. Nevertheless, with the recognition of more and more examples of natural substances, as innocent as lactose, that are able to increase the risk of cancer [phaeochromocytoma in rats (Roe and Baer, 1985)], with increasing recognition of the extent of DNA damage by endogenous electrophiles and with improving understanding of the normally high rate of DNA repair, only obligate ostriches still have their heads buried in the sand.

How, then, should we test for carcinogenicity?

A serious problem with carcinogenicity tests in rodents has been high incidence of so-called 'spontaneous' benign and malignant neoplasms in untreated control groups. Curiously, although the fundamental aim of carcinogenicity tests in rodents has been to discover the causes of 'spontaneous' cancers in man, there has been remarkably little interest in the causation of the tumours that arise in untreated groups!

Worse still, although there is now overwhelming evidence that caloric intake is a major determinant of cancer incidence in rodents not deliberately exposed

to any known carcinogen or xenobiotic agent, tests for carcinogenicity are still carried out under conditions of life-long overnutrition and under conditions in which exposed and control animals are not necessarily comparable in terms of food intake. Failing to compare 'like' with 'like' in this way constitutes what epidemiologists would describe as 'ignoring an important confounding variable'.

I have yet to encounter what I would regard as a valid argument against the routine use of mild dietary restriction as a standard procedure in all carcinogenicity tests. I am not impressed by the argument that if one introduced dietary restriction into such testing one would no longer be able to use databases built up from studies on animals fed *ad lib*. Such databases are in my view largely valueless anyway, since they refer to animals rendered unphysiological through overfeeding. In rats overfeeding has, *inter alia*, effects on the kidneys (chronic progressive nephropathy), the heart (chronic myopathy), blood vessels (polyarteritis), the liver (increased absolute and relative liver weight with perpetual *P*-450 enzyme induction), the pituitary (hyperplasia and neoplasia, particularly of prolactin-producing cells), mammary glands (hyperplasia, increased secretory activity, benign and malignant neoplasia), the pancreas (islet cell neoplasia), the adrenal medulla (focal hyperplasia and neoplasia in some strains) mineral metabolism (parathyroid hyperplasia and neoplasia, pelvic nephrocalcinosis, metastatic calcification in the aorta, glandular stomach mucosa, gut mucosa, lungs, kidneys and many other tissues), the spinal cord and corda equina (radiculo-neuropathy), and an increased risk of benign and malignant neoplasia at virtually all body sites. I can think of no valid argument for preferring the use of animals at high risk of developing any or all these conditions to the use of physiologically and pathologically normal and healthy animals.

That is the first and most important point I wish to make with regard to how carcinogenicity tests should be conducted. The second point arises out of the fact that, nowadays, a majority of carcinogenicity tests in rodents are conducted on substances that have given negative or, at worst, equivocal results in tests for genotoxicity. This means that the main object of carcinogenicity tests is to see whether test agents can act indirectly as carcinogens (i.e. as non-genotoxic carcinogens). A feature of non-genotoxic carcinogenic activity is that cancers usually or always arise out of a background of disturbed physiological status and/or of evidence of hormone-mediated or regenerative hyperplasia. This being so, there is a strong case for paying far more attention than in the past to non-neoplastic pathology and particularly to measures of hormonal status, mineral balance and regulatory peptide activity during studies. At present the long list of measurements made, mainly in satellite groups of animals in carcinogenicity studies, is determined far more by the availability of test methods than by the likely value of information gained from them. Furthermore, little or no effort is made to integrate the findings into any composite picture. In particular, no effort is made to study the various effects of exposure to test agents in individual animals. Instead, means for various parameters are reported for whole groups of animals with the result that the chances of observing patterns of response affecting a range of different parameters in individual animals are greatly reduced. All this surely needs to change and to change radically. In particular, new micro-methods are urgently needed to monitor the physiological and hormonal status of individual animals throughout the duration of carcinogenicity tests.

Certainly, those of us who have many years of experience in carcinogenicity testing have come a long way since the days of simple lump-counting, but there is still a great deal of scope for increasing the value of such tests.

[Francis J. C. Roe—Consultant in Toxicology]

REFERENCES

Ames B. N. (1989) Mutagenesis and carcinogenesis: endogenous and exogenous factors. *Environmental Mutagenesis* **13,** 1–12.

Ames B. N. and Gold L. S. (1990) Too many rodent carcinogens: mitogenesis increases mutagenesis. *Science* **24,** 970–971.

Ames B. N. and Saul R. L. (1987) Oxidative DNA damage. *Annals of Internal Medicine* **107,** 526–545.

Armitage P. and Doll R. (1961) Stochastic models for carcinogenesis. In *Proceedings Fourth Berkeley Symposium on Mathematic Statistics and Probability. IV. Biology and Problems of Health.* p. 19.

Berenblum I. and Shubik P. (1947) A new quantitative approach to the study of the stages of chemical carcinogenesis in the mouse's skin. *British Journal of Cancer* **1,** 8–15.

Borghoff S. J., Andersen M. E. and Conolly R. B. (1991) Protein nephropathy and kidney cancer of male rats: qualitative and quantitative issues of human relevance. *CIIT Activities* **11** (1), 1–6.

Borghoff S. J., Short B. G. and Swenberg J. A. (1990) Biochemical mechanisms and pathobiology of $\alpha_{2\mu}$-globulin nephropathy. *Annual Review of Pharmacology and Toxicology* **30,** 349–367.

Cohen S. M. and Ellwein L. B. (1990) Cell proliferation in carcinogenesis. *Science* **249,** 1007–1011.

Heasman M. A. and Lipworth L. (1966) Accuracy of certification of cause of death. *Studies in Medical and Population Subjects No.* 20. General Register Office, HMSO, London.

Iversen O. H. (1988) Initiation, promotion: critical remarks on the two-stage theory. In *Theories of Carcinogenesis.* Edited by O. H. Iversen. pp. 119–126. Hemisphere, Washington.

McGuinness E. E., Morgan R. G. H., Levison D. A., Frape D. L., Hopwood D. and Wormsley K. G. (1980) The effects of long-term feeding of soya flour on the rat pancreas. *Scandinavian Journal of Gastroenterology* **15,** 497–502.

Peto R., Roe F. J. C., Lee P. N., Levy L. and Clack J. (1975) Cancer and ageing in mice and men. *British Journal of Cancer* **32,** 411–426.

Report of Joint Working Party of the Royal Colleges of Pathologists, Physicians and Surgeons. (1991) *The Autopsy and Audit*. Royal College of Pathologists. pp. 1–24.

Roe F. J. C. (1988) Let us forget "tumour promotion" and start thinking about mechanisms of "non-genotoxic carcinogenicity". *APMIS* **96** (Suppl. 2), 91–99.

Roe F. J. C. (1989) Non genotoxic carcinogenesis: implications for testing and extrapolation to man. *Mutagenesis* **4**, 407–411.

Roe F. J. C. (1991) 1200-Rat Biosure Study: Design and overview of results. In *Biological Effects of Dietary Restriction*. Edited by L. Fishbein, pp. 287–304. ILSI Monographs: Springer–Verlag, New York.

Roe F. J. C. and Baer A. (1985) Enzootic and epizootic adrenal medullary proliferative disease of rats: Influence of dietary factors which affect calcium absorption. *Human Toxicology* **4**, 27–52.

Roe F. J. C. and Lee P. N. (1991) Carcinogenicity tests. *Lancet* **337**, 587.

Roe F. J. C., Lee P. N., Conybeare G., Tobin G., Kelly D., Prentice D. and Matter B. (1991) Risks of premature death and cancer predicted by bodyweight in early life. *Human Experimental Toxicology* **10**, 2185–2288.

Roe F. J. C. and Wood D. (1992) Acetaldehyde and formaldehyde: is there a cancer risk for man? *Indoor Environment* **1**, 8–15.

Rous P. and Kidd J. G. (1941) Conditional neoplasms and sub-threshold neoplastic states. *Journal of Experimental Medicine* **73**, 365–390.

Weinstein I. B. (1991) Mitogenesis is only one factor in carcinogenesis. *Science* **251**, 387–388.