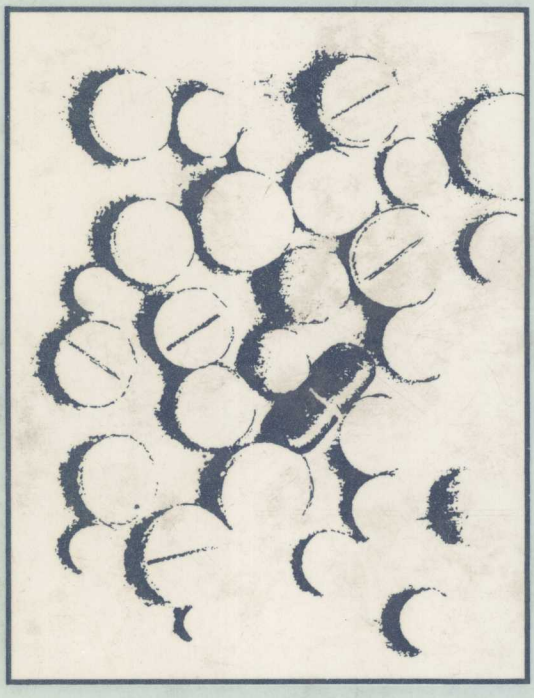
Risk–benefit analysis in drug research



Edited by J. F. Cavalla

Risk–Benefit Analysis in Drug Research

Edited by

J. F. Cavalla
Wyeth Laboratories, Taplow, Maidenhead, Berkshire, England

Proceedings of an International Symposium
held at the University of Kent at Canterbury, England,
27 March 1980

MTP PRESS LIMITED
International Medical Publishers
LANCASTER · BOSTON · THE HAGUE

Published by
MTP Press Limited
Falcon House
Lancaster, England

First published 1981

ISBN 0-85200-364-1

Printed in Great Britain by
Mather Bros (Printers) Limited, Preston

Risk–Benefit Analysis in Drug Research

List of Contributors

Sir Douglas Black
Royal College of Physicians,
11 St Andrew's Place,
Regent's Park,
London NW1 4LE, England

Professor A. M. Breckenridge
Department of Pharmacology and Therapeutics,
University of Liverpool,
Liverpool L69 3BX, England

Dr R. W. Brimblecombe
Vice-President, Research and Development,
Smith, Kline and French Laboratories,
Welwyn Garden City,
Herts., England

Dr J. F. Cavalla
Research Director,
Wyeth Laboratories,
Taplow, Maidenhead,
Berks. SL6 0PH, England

Dr C. N. Christensen
Vice-President,
Lilly Research Laboratories,
Indianapolis,
Indiana, USA

Dr A. D. Dayan
Wellcome Research Laboratories,
Beckenham,
Kent BR3 3BS, England

Professor C. T. Dollery
Royal Postgraduate Medical School,
Hammersmith Hospital,
Ducane Road,
Hammersmith,
London W12 0HS, England

Dr J. F. Dunne
Senior Medical Officer, Pharmaceuticals,
World Health Organization,
1211 Geneva 27, Switzerland

Rt Hon. David Ennals, MP
Former Secretary of State for Social Services,
The House of Commons,
London SW1A 0AA, England

Professor D. G. Grahame-Smith
Rhodes Professor of Clinical Pharmacology,
University Department of Clinical Pharmacology,
University of Oxford,
Radcliffe Infirmary,
Oxford OX2 6HE, England

Dr W. H. W. Inman
Director, Drug Surveillance Research Unit,
University of Southampton,
Southampton SO2 3FL, England

Dr P. A. J. Janssen
Janssen Pharmaceutica,
B-2340 Beerse, Belgium

Dr S. Lock
Editor, *British Medical Journal*,
Tavistock Square,
London WC1H 9JR, England

Dr J. Maddox
Editor, *Nature*,
4 Little Essex Street,
London WC2R 3LF, England

Dr B. B. Newbould
Research Director,
ICI Ltd: Pharmaceuticals Division,
Alderley Park, Macclesfield,
Cheshire SK10 4TG, England

Sir Edward E. Pochin
National Radiological Protection Board,
Harwell, Didcot,
Oxon., OX11 0RQ, England

Mr J. D. Spink
Regulatory Controller,
The Wellcome Foundation Ltd,
187 Euston Road,
London NW1 2BP, England

Preface

The appreciation of risk like the awareness of beauty lies very much in the eyes of the beholder. It involves a value judgement and can never be absolute. Yet paradoxically, modern society is demanding ever greater degrees of safety in the medicines it takes, to the extent that nothing short of the total absence of risk will be tolerated.

Since 1960, and mainly as a result of the thalidomide tragedy, governmental regulation of testing and use of new medicines has grown apace throughout the world. It has derived impetus not only from the understandable wish of the public to seek protection, but also from the anxiety of bureaucrats and politicians not to be seen to have made mistakes. Both these concerns have been inflamed by the recognition of the media that all drugs make news and horror drugs make the best news of all.

Prior to this time the physician and his cures enjoyed a relatively supportive public. It was true that quacks existed and were recognized as such but, in the main, people wanted to take medicines and expected them to do them good. Side-effects were little recognized and when observed were often attributed to the disease rather than the cure. Only as education grew and communication improved did it begin to be appreciated that medicines might cause disease as well as prevent it. With the concomitant growth of the pharmaceutical industry in the second half of this century and the consequent introduction of many potent new medicines, the concern for safety became overriding; to the extent that many traditional remedies were dropped for fear of their side-effects.

While in most cases the concern for safety is a direct humanitarian response embodying the desire of all to prevent suffering, on occasions it is tainted by venal motives. No one can be seen publicly to support hazard, though cigarettes are still advertised and highwire acrobats applauded. When medicines are being considered, however, an emotional response seems to be elicited and rational discussion obscured. Moreover, special pleading either for commercial gain in the form of financial compensation for damage done or political advantage in seeking public control of an essentially entrepreneurial industry can often be detected. All these factors tend to militate against

rational discussion of an essentially pragmatic problem. Evermore the quest for absolute safety and abolition of all risk in medicines prevents the assessment of the comparative risk–benefit equation.

This book records an attempt by several parties to determine how this paradox can be resolved: how best to equate the risk involved in taking a new medicine with the potential benefit it can bestow. It is timely such an effort was made. The international pharmaceutical industry is responsible for most, if not all, the new medicines now being offered to man. It supports an extensive and growing research effort throughout the world to undertake this task yet, perversely, as the expenditure has increased so the discoveries have lessened. With the growth in cost and the decline in invention, smaller companies have moved out of ethical research. In the United Kingdom alone, a dozen research units have closed in the last 15 years as a result either of mergers or business failure. In the view of many in the industry the time has now come when further demands for safety coupled with ever greater regulatory control might well jeopardize the future of the enterprise. In effect it will become so difficult and so onerous a task to introduce a new medicine that no one will be prepared to take it on. Already it is costing in excess of $20 million and taking over 12 years to bring a new drug to the marketplace. Only substantial reversal of the trend can bring effective change.

It was in this climate that the Society for Drug Research decided to arrange a meeting to discuss the whole question of risk–benefit analysis in drug research. Invited speakers included politicians, members of government regulatory bodies, representatives of the media, physicians, clinical pharmacologists, toxicologists and research directors. At no time was there an adversarial approach to what was said. All accepted the need to discuss a serious problem with the aim of achieving a consensus for resolution.

Clearly surcease in regulatory control is necessary if drug research as we know it is to survive. For this, responsible leadership is required to educate the public to accept the fact that medicines like surgery must always possess some small element of risk if their benefit is to be made manifest.

J. F. CAVALLA

Contents

1
Risk–benefit in medicine

SIR EDWARD E. POCHIN

Perhaps the first thing to say about risk–benefit analysis is that everybody does it every day; and the second is that it usually cannot really be done, at least in any quantitative way. We are, in fact, surely making some kind of risk–benefit evaluation every time we choose between alternatives: we are weighing the advantages against the disadvantages. Our assessment may be heavily influenced by habits, by traditions, by a few misconceptions and by yesterday's headlines, but some sort of judgement will be made. To do this job properly however, and to strike a true balance between risk and benefit, we need two sorts of information that we usually have not got. We need a factual and, essentially, a numerical estimate of all components of the risk, as well as of the benefit, if we are to dignify the process by the name of risk–benefit analysis. Much more important and much more difficult, however, we need to have some idea of how much weight should be put on the different components of the risk and of the benefit; and we need to assess numerically the relative importance of these various components. We may say to ourselves that the risk is a small one, and not much of a worry, or we may call it a small risk but an unpleasant one that we do not like to take, or we may recognize that the benefit is trivial, but we want it. Our personal weighting factors are often more important than the bare arithmetic of profit and loss.

To tackle the problem of risk–benefit analysis it is essential to be numerical about the size of the risk, as well as to be concerned with the weighting that should be put, for example, on the safety of the drug or its value in treatment, and the weighting that other people will put on these factors. Because inevitably any tidy-minded risk–benefit equation will come out in apples and pears—in factors that are not easily expressed in commensurable terms.

This is particularly true in medicine and in public health problems, where one will be dealing with deaths, with non-fatal diseases, and with non-fatal disabilities to which relative weighting factors must be attached; and one must

be aware also of the anxieties and stresses about the possibility of those deaths, diseases and disabilities.

Obviously this is a very large and awkward problem, even when we take the easy way out, of looking only at numbers of deaths on both sides of the equation, and ignoring, or putting less weight on, the non-fatal components. Even then it is not easy. For example, what about advising someone to have an operation that carries a 2% risk of immediate death at the time of the operation but that offers a 10% chance of avoiding death from fatal malignant disease two years later? Or, is it worse to cause one death by vaccinating a healthy child, or to allow one death to occur through having failed to prevent the disease by doing so? We are comparing one death with one death, but they are not equal deaths, and the comparison is made no easier by the risks of vaccination being so low. There were about three deaths per year between 1967 and 1976 from the effects of vaccination (England and Wales[1], categories E933 and 934), during a period in which an average of 4 million vaccinations were carried out each year (against the seven diseases listed in 'Social Trends'[2]), indicating a risk of fatality in the region of one in a million such vaccinations.

How can we visualize what 10^{-6} means as a risk? We can obtain some perspective by a comparison with other circumstances giving rise to the same risk, for example, that cigarettes appear to carry a risk of causing death of a little over 10^{-6} per cigarette smoked[3], any such death involving an average of about 5 years' loss of life expectancy. Numerically, therefore, vaccinations may have had a risk equal to that from one cigarette in terms of deaths attributable to each; or to a packet of cigarettes if the risk were expressed as the average years of life expectancy lost per vaccination.

However difficult it may be to appreciate the importance of different types of risk, this appreciation can only be helped by assessing numerically the magnitude of the risks or the frequency of the benefits, so that decisions can be influenced, at least in part, by a factual knowledge of their likely consequences. Certainly we cannot add apples and pears, but we may know that we would rather have two pears than five apples.

It is useful therefore, that this habit of risk watching seems to have proliferated so much in the past few years[4,5]. It does not answer questions, but it poses the questions in a more exact form and can help one to guide one's decisions in the light of the facts. It is a curiously obsessive habit. It is almost like collecting stamps or engine numbers, and in a way it is about as juvenile. But at least numbers are produced, and actions can be rated according to the likelihood of certain specified consequences, particularly of harmful biological effects. This, I believe, has a certain value, as long as no-one believes that it will get him all the way to answering the question.

Even when this is done, one is obviously only looking at the findings in a particular problem under particular circumstances, as averaged over a certain period of time, and one can only say that if circumstances do not change, then

it is to be expected that the risk will not change, and the risk estimate, of so many effects per million exposed to a particular challenge, can be used in forecasting. One point that has to be clearly remembered, however, is that in events having a low risk, for example of death, the number of deaths will be very small unless very large populations are considered. One has to be very much aware therefore of the confidence limits of the risk estimates that emerge. For example, if we see[6] that in 23 000 liver biopsies there were four deaths, that figure of four has 90% confidence limits from 1.4 to 9.2 on ordinary Poisson statistics, and so one must be fairly wary about the accuracy of any rate derived.

Also, obviously, one has to be wary about the way in which the estimate lumps together a variety of different levels of hazard. I have quoted the liver biopsy data as an example. This hazard of liver biopsy will vary in people of different ages. It certainly varies in groups with different severity of liver disease. It will be considerably affected by the likelihood of bleeding, so will vary with the blood coagulability and the degree of portal hypertension and doubtless with many other factors. The cake may be split into as many sections as one wishes, with finer and finer subdivisions, until finally it is all crumbs and there is nothing profitable to use. Any good statistician can claim that every average is nonsense because it could have been split a little bit more finely, or a lot more finely. But, if one assays for a particular population and can hope for reasonable homogeneity within the population, one can get a figure that applies as an average to a similar population; and this can help.

There is one obvious example in which this sort of analysis can help, and has helped, in the work that has been done on X-ray screening for cancer of the breast, in a number of countries. Obviously the survival rate of patients with cancer of the breast can be increased, or we hope it can be increased, by the earlier diagnosis that is obtained by X-ray screening of the healthy population. On the other hand, the irradiation of the breast may itself induce a certain number of cancers, and the task is to find that number. There is thus a risk–benefit problem in the screening of healthy populations for cancer of the breast and we need to know which wins. Does the screening programme induce more cancers than it saves, or save more than it induces? More specifically, we want to know above what age to screen, and below what age we shall be worse off by screening because of the irradiation. This is a situation in which the induction rate of breast cancer from radiation at a given dose is quite reasonably reliable[7]. There is good evidence on the rate of induction at low dose. That is simply because data have been obtained on studies by fluoroscopy in the course of pneumothorax treatment, and each individual exposure was at reasonably low dose. One can get, therefore, a good figure for the induction at the appropriate dose of X-rays, or only somewhat higher, and there are reasonably good data on the age variation of this induction rate. Obviously the rate of induction will vary with the type of examination, because the mean X-ray dose to breast will vary. The mortality of the breast

tumours that may be induced by radiation must be known, as must the time interval between irradiation and detection of the cancer or death from it. These latencies for radiation-induced malignancies are long, probably of the order of 20, 25 or more years median interval between irradiation and even the detection of an induced malignancy. Thus when one is reviewing the age at which to perform such examinations, one is concerned with this figure of latency; and one will of course need to know also the natural incidence and mortality at different ages without a screening programme.

Studies of this sort have been done in a number of countries, and the most reliable and extensive suggest that above the age of 50 in women we are winning; more cancers will be detected than will be induced. Obviously examinations at greater ages will carry less penalty because of the long latency, if induced cancers are not expressed during the remaining lifetime. Obviously also more will be detected by such examination at these ages because the incidence is rising rapidly with age. Above this critical age of about 50, therefore, it seems likely that more will be detected and prevented than are induced[8], and below the age of 50 it may be the reverse.

There is particularly good work from Japan on screening for stomach cancer[9]. This malignancy is responsible for the greater part of all cancer mortality in Japan. It is also an important cancer to try to detect by mass survey of healthy people, because it is one which is very liable to be fatal, unless detected before symptoms have started to develop. Also, X-ray examination is really the only practicable method of screening on a mass scale, whereas for the breast there are alternatives.

The quantitative basis that has been used in Japan is very much helped in its precision by the fact that, because of the commonness of stomach cancer there, screening was started in 1960, and as many as 4 million people were screened in this way during 1978. Iinuma and his colleagues, therefore, had good evidence as to the frequency with which otherwise undetected stomach cancers were found by the technique of screening that they used. They had good evidence on the frequency with which the cancers so detected were cured, or how frequently they caused death despite their being so detected and they had the data on the variation of all those figures with age and sex. They obtained information on the radiation dose, from the examinations that were made, to all the body organs within the radiation beam; and there are now good data on the induction rate of malignancies for most of the main body organs by radiation[10,11]. It was, therefore, possible to estimate with some confidence in this survey the number of cancers that might be detected and the number induced, and make a comparison with the natural incidence rate and the mortality without screening.

Since he was concerned with age variation, Iinuma compared the two situations, with screening or without, in terms of the years of life expectancy lost. Table 1.1 shows for males the years of life lost with screening because of cancers induced, and these drop rapidly with age at examination. Taking

account also of the mortality from cancers that caused death in spite of the screening programme, an estimate is obtained of the total loss of life in years per thousand person-years of screening. Without screening and with only a small average radiation dose to the population, by X-rays given to patients who came up because of symptoms, cancer induction forms a trivial component of the risk but the later diagnosis of the naturally occurring cancers gave substantial estimates of life loss.

Table 1.1 Radiological screening for gastric cancer in Japan, showing years of life lost per 1000 men from induced and gastric cancer

	From Cancers	*Age*		
		25–30	*30–35*	*35–40*
With screening				
Greater average radiation	induced	1.20	0.92	0.67
Earlier diagnosis	gastric	0.44	0.87	1.38
	total	1.64	1.79	2.05
Without screening				
Less average radiation	induced	0.01	0.01	0.01
Later diagnosis	gastric	1.12	2.23	3.52
	total	1.13	2.24	3.53
Net benefit from screening; years (per 1000)		−0.51	+0.45	+1.48

(Iinuma, *et al.*)

Thus, screening of the populations between age 35 and 40 shows a gain of 1.5 years per thousand, at ages 30 to 35 the gain is 0.5 years, but screening at age 25 to 30 results in a loss of life expectancy. The total risks and benefits balance at an age of about 30, both for males and for females, above which there is a gain by preventing more cancers than are induced, and below which there is a loss by inducing more than are prevented in the younger age groups. Quite obviously there are a lot of uncertainties in this analysis, but, given the assumptions that were made, and the care with which it was done and the large number of factors taken into account, it does indicate that for the Japanese population, screening is likely to be profitable in saving life above the age of 30. These studies of screening programmes are important particularly because the programmes may involve large numbers of people and therefore appreciable numbers of cancers might be induced by radiation; and other risks might be involved in other types of screening procedure.

There are similar problems, in some ways more difficult, in looking at the use of various procedures in medical or clinical research, procedures such as studies by X-rays and radio-pharmaceuticals. There is no reason to apologize for giving examples in which radiation risk analysis is involved. Radiation

hazards have been very much better studied and documented than very many environmental chemical hazards, and many pharmacological hazards. There has been a great deal of work on the epidemiology of radiation induction of malignancies and of genetic effects and estimates are now reasonably reliable[10].

Table 1.2 The radiation exposure levels of patients undergoing various diagnostic tests using X-rays and radionuclides

Equivalent whole-body dose (mrem)	*Number of types of diagnostic tests using:*	
	X-rays	*Radionuclides*
0.01	—	1
0.1	—	3
1	1	21
10	18	33
100	14	25
1000	—	5
10 000	—	—
Geometrical mean dose	74 mrem	37 mrem
Risk	8×10^{-6}	4×10^{-6}

In medical research, for example using radiopharmaceuticals, no problems need to arise if the test is being done for clinical purposes in any case, and it is merely its results that are being studied. The position is usually similar even if somewhat larger activities are used, or rather larger samples of blood are removed for the purposes of research. But, if healthy or other volunteers are being examined, who cannot be expected to benefit from the tests, then it is important to look at the risk–benefit balance and ensure that the risk to the individual is trivial, and that the size of the benefit certainly outweighs the risk.

But how small is the risk likely to be? Table 1.2 shows, for the range of conventional X-rays, the amount of radiation exposure involved. For a group of 30 or so typical diagnostic procedures the geometric mean dose is of about 70 mrem. (This value is the 'effective' whole body dose, either from radiation of the whole body uniformly, or as the whole body dose equivalent in number of harmful effects to doses delivered to individual body organs.) Conventional radionuclide investigations give a rather lower figure. The risks of inducing any fatal effect implied by these geometric mean values are of about 8×10^{-6} for X-rays, and 4×10^{-6} from radionuclide examinations. For perspective on the mean doses, the whole body receives about 100 mrem per year from natural sources and has always done so. As perspective on the size of the risk (although not necessarily on its acceptability) the smoking of six, or three cigarettes corresponds in risk of fatal effect to the 8 or 4×10^{-6} mean risks of typical examinations by X-rays or radiopharmaceuticals.

How far is this simply a theoretical calculation? Three studies have been made on large numbers of patients after particular radiological procedures in

which there has been a detectable cancer incidence above expectation (Table 1.3). Following multiple fluoroscopies, when the patient was facing the tube so that the breast received a somewhat higher dose than usual, the breast cancer induction rate corresponded to a mortality of about 10^{-4} per examination[12]. The use of diagnostic X-rays involving the pelvis of a mother during pregnancy was followed by a small excess of cancers in the fetus, as expressed during the first 10 years of life, in this case with a mortality[13] of about 2×10^{-4}. Studies of treatment for ringworm involving a few rads to the thyroid in about 11 000 children showed a detectable excess of thyroid cancer. The mortality of radiation-induced cancers of the gland is low so that again there is likely to have been a fatality risk of about 10^{-4} per treatment[14].

Table 1.3 Radiological procedures involving a detectable cancer incidence above expectation

Procedure	*Type*	*Cancer incidence*	*Mortality*
Fluoroscopy (facing tube)	Breast	2.10^{-4}	1.10^{-4}
In utero, pelvic X-ray	All types	2.10^{-4}	2.10^{-4}
Thyroid dose, 7 rad	Thyroid	9.10^{-4}	1.10^{-4}
Thorotrast (25 ml)	Leukaemia	120.10^{-4}	
	Liver	500.10^{-4}	

We are seeing X-rays and radiopharmaceuticals therefore with typical estimated risks of a few deaths per million, and evidence of certain such procedures with observed risks which are factors of 1.5 or 2.0 orders of magnitude higher than this. (The use of Thorotrast, the thorium oxide contrast medium, involved very much higher risks, in the order of 6×10^{-2} per examination.)

In view of this range of hazards that can be estimated for radiopharmaceuticals or X-rays, the World Health Organisation (WHO), some 3 years ago, produced a report[15] on the use of investigations involving radiation in research on man, as well as in other circumstances. They divided the diagnostic tests that could be used with radiation into three categories (Table 1.4).

The boundaries between WHO's categories correspond to 50 mrem, 500 mrem and 5 rem whole body doses. The average risk of causing a fatal malignancy in a man of 40 can be estimated to be in the region of 4×10^{-6} at the boundary between categories I and II, this estimate varying in proportion to dose in the higher categories[16].

How can one illustrate the size of these risks? The average loss of life expectancy would be small, varying from less than half an hour from the effects of category I, to 1 or 2 days in category III, if it was justifiable to average the duration of life loss from an induced cancer amongst all recipients of the test. Or, on a more reasonable comparison, the radiation risk of fatality at the