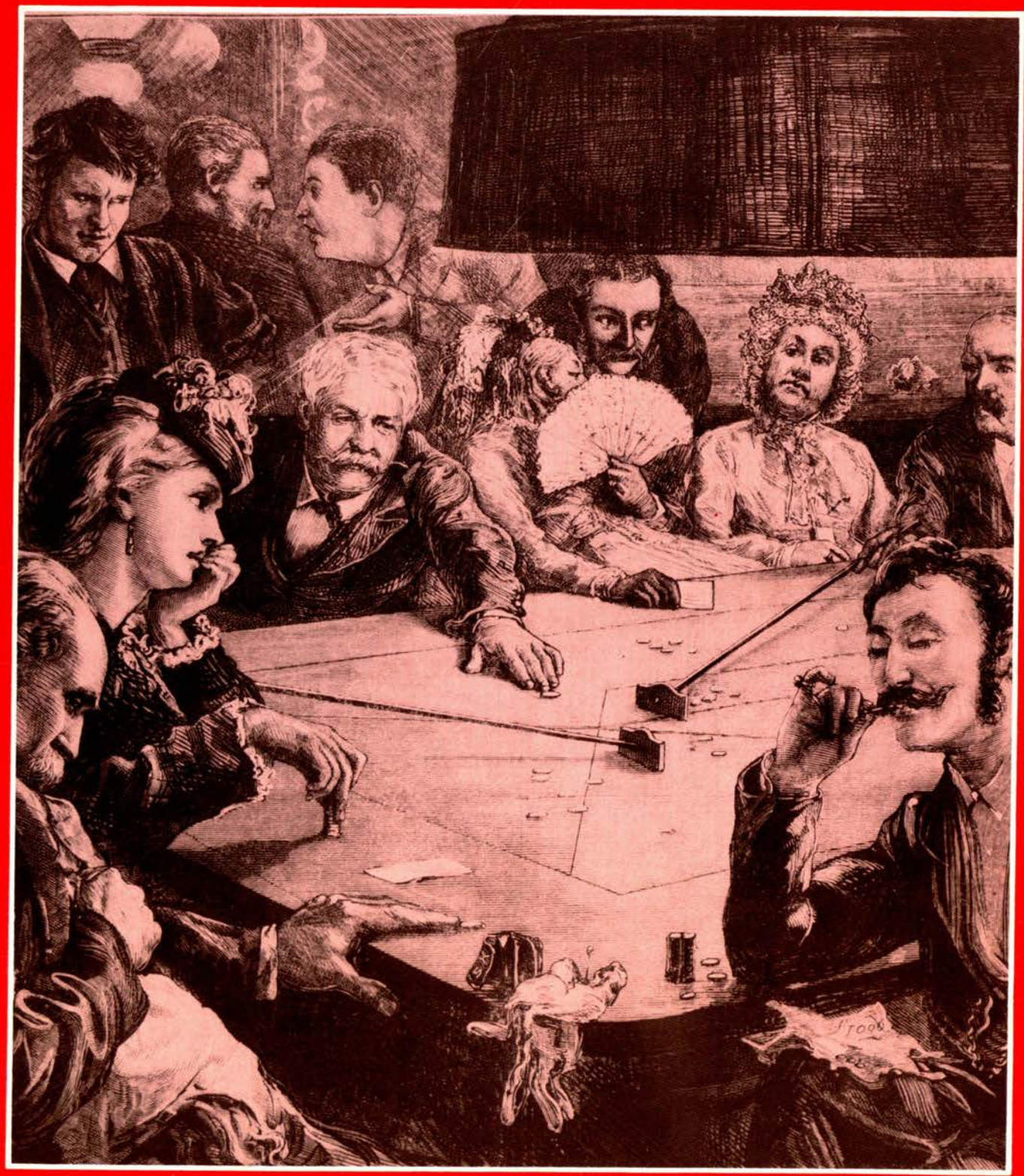
'WHAT ARE MY CHANCES DOCTOR?'

A REVIEW OF CLINICAL RISKS



'WHAT ARE MY CHANCES DOCTOR?' – A REVIEW OF CLINICAL RISKS

Drawing by S. Harris. Reprinted with permission of The New Yorker Magazine, Inc.

Bernie O'Brien
Health Economics Research Group
Brunel University

November 1986

ACKNOWLEDGEMENTS

My thanks to George Teeling Smith for his continued enthusiasm, guidance and patience on this project. Also thanks to Mary Elliott and Karen Lane for their help and unerring ability to read my handwriting. The errors that remain are obviously my responsibility alone.

BERNIE O'BRIEN

Office of Health Economics
12 Whitehall London SW1A 2DY

No 84 in a series of papers on current health problems published by the Office of Health Economics.
For previous papers see page 44.

ISBN 0473 8837

Price £1.50

Office of Health Economics

The Office of Health Economics was founded in 1962 by the Association of the British Pharmaceutical Industry. Its terms of reference are:

To undertake research on the economic aspects of medical care.

To investigate other health and social problems.

To collect data from other countries.

To publish results, data and conclusions relevant to the above.

The Office of Health Economics welcomes financial support and discussions on research problems with any persons or bodies interested in its work.

CONTENTS

FOREWORD

Half a century ago little attention was paid to the risks associated with medical and surgical treatment. The hazards of sickness itself were so obvious, that the considerable risks of medical intervention were more or less taken for granted.

In 1986 the situation is completely different. The classical 'killer' diseases affecting children and young adults have largely been eliminated as a result of therapeutic progress. Much of medical care is now concerned with prolonging life in cases of chronic disease such as diabetes or hypertension or alternatively with improving the quality of life for people with diseases like arthritis or schizophrenia.

In addition, 'high-tech' medicine is replacing the simple procedures of the 1930s, and some of the recent innovations are associated with significant new risks.

Thus doctors are now more than ever faced with the problem of balancing immediate risks against long-term benefits. And health providers as a whole, including those who have discovered and manufactured modern medicines, are consequently sometimes accused of exposing patients to unjustifiable hazards.

This paper by Bernie O'Brien of the Health Economics Research Group at Brunel University is part of a continuing attempt to put these problems into perspective. It should improve its readers' understanding of the difficult exercise of balancing the risks associated with disease against the often much publicised risks associated with their treatment.

As Bernie O'Brien points out 'risk' is a relatively new and largely professional concept. Patients generally think in terms of their 'chances' both of recovery and of possible harm from their treatment. The health professions, the consumerist organisations and the media need to develop a more sophisticated awareness of the ways in which patients can be helped to assess the relative risks of their illness and their treatment. This paper should help to provide this all-important awareness. That, in turn, should enable doctors and their patients to make more rational decisions about the risks and benefits of modern therapies.

GEORGE TEELING SMITH

'Risks are among the facts of life. In whatever we do and whatever we refrain from doing, we are accepting risk. Some risks are obvious, some are unsuspected and some we conceal from ourselves. But risks are universally accepted, whether willingly or unwillingly, whether consciously or not.'

Pochin (1975)

'Somewhere between 1910 and 1912 in this country . . . a random patient, with a random disease, consulting a doctor chosen at random had for the first time in the history of mankind, a better than fifty-fifty chance of profiting from the encounter.'

Henderson (1977)

INTRODUCTION

Clinical practice, be it medicine or surgery, involves calculated risks. Health can only usually be restored or conferred at the price of accepting the risks of treatment. Medicine is a gamble which is expected to pay off.

Consider the patient suffering from coronary heart disease and the pain of severe angina. One of the treatment options available to him is surgery – coronary artery bypass grafting. But with this treatment there is about a 1 in 30 chance that he will be dead within thirty days of the operation (English *et al*, 1984). The question of whether he would accept surgery will depend upon how he trades-off the fatality risk against the pain. It will depend upon his *attitude* to risk, which will depend upon how he *perceives* the risks. In turn, this will depend upon how much risk information the doctor gives him and in what format.

The aim of this paper is to examine the nature of such treatment gambles and the factors that determine how risks are estimated, perceived and evaluated in the context of clinical decision analysis. The question of patient preferences in such choices is discussed in detail with special emphasis on patient attitudes to risk as a factor influencing decision making. The paper is essentially a review of clinical risks and clinical risk-taking.

In **section 1**, some semantic 'ground clearing' is undertaken where the various uses of the word risk are examined. Building on a common definition of risk, the nature, meaning and estimation of probability is discussed in **section 2** and serves as a preface for the exposition of the principles of expected outcome, value and utility in **section 3**. The use of probability in the evaluation of health states using the 'standard gamble' is explored in **section 4**.

The discussion of the general problems of risk assessment in **section 5** is a preface to the analysis of how risks are assessed in prescription medicines and surgery in **section 6**. The evidence on risk perception and misperception by patients and doctors is reviewed in **section 7** and the question of whether medicine is a voluntary or an involuntary risk is considered in the context of legal definitions of informed patient consent with respect to treatment risks in **section 8**. Finally in **section 9** the various aspects of the paper are brought together in a discussion of the evaluation of clinical risk taking.

1 RISK: A WORD OF UNCERTAIN MEANING?

The word 'risk' originates from the French *risqué* and is quite a modern addition to the English language. Collinson and Dowie (1980) report that the word did not enter the language until the mid-seventeenth century and then appeared in anglicised spelling in insurance transactions during the second quarter of the eighteenth century. Only in the twentieth century did the derivative risque come into general use to describe something daring or salacious.

One of the immediate difficulties that arises when trying to define risk is that it has both a set of technical definitions – risk assessment, hazard management etc and a strong colloquial tradition. In common usage the word risk is intrinsic with the chances of loss rather than gain. The Oxford English Dictionary definition of risk includes '. . . hazard, chance of bad consequences, loss etc'. Yet in the analysis of decision-making in medicine and other areas, the potential costs of a decision are only one side of the coin; individuals will accept risks in order to gain some benefit. As Shakespeare (1594) stated:

> '. . . men that hazard all
> Do it in hope of fair advantage.'

This poetic rendition of decision-making under uncertainty illustrates the nature of the risk-benefit trade-offs that we all make every day of our lives. The consideration of risk in decision-making is to recognise that the future can never be known with certainty and, at best, comes with a range of probabilities attached to it. The reason we accept risks – be they travel risks, occupational risks or medical risks – is in order to gain some benefit. But note that Shakespeare's benefit measure '. . . the hope of fair advantage' is less than certain; the gains from a decision being probablistic as well as the costs (risks).

Risk and uncertainty are often used interchangably. Both concepts can be defined as being basically a problem of lack of *information* about future 'states of the world' that might arise. Knight (1933) distinguished between the two concepts by arguing that **RISKS** are future outcomes to which it is possible to attach probabilities, whereas **UNCERTAINTY** (pure) is a situation where the individual cannot attach probabilities to future 'states of the world'. This paper concerns itself with risks, where probabilities – be they objective or subjective – can be attached to outcomes.

Semantic confusion commonly surrounds the discussion of risk in relation to **SAFETY**. Often safety is simply taken to be the inverse of risk. A number of commentators have preferred to use the terminology of safety in preference to risk. Green and Brown (1978), essentially argue that risk and safety lie at opposite ends of the same spectrum. Furthermore, public agencies involved with risk regulation tend to prefer the safety lable – eg Health and Safety Executive, Committee on Safety of Medicines. But as Green and Brown (1978) recognise, safety is a word which is open to abuse because it implies some absolute standard – 'safe' – is an obtainable objective.

But complete or absolute safety (zero risk) is an extreme situation where the probability of an event occurring is strictly zero. A moment's reflection indicates that this state can never be achieved because all future events and states of the world cannot be observed or predicted with certainty. Strictly speaking, nothing is impossible and therefore nothing is *absolutely* safe. (In 1986, the Chernobyl nuclear power plant disaster and the Space Shuttle disaster amply illustrate this point).

Therefore in its typical usage the label 'safe' is used to reference some level of risk which has been *judged* to be so small that it is insignificant and equivalent to zero; the idea of an *acceptable* risk. If the 'safe' label is to be used we must recognise that it is the result of some evaluative process. Generally, risk is more usefully viewed as continuous variable, rather than safety which lends itself too easily to the misleading safe/unsafe dichotomy.

The Royal Society (1983) took great pains to disentangle the semantics of risk, examining the more common synonyms and their interpretation. One of the terminology guidelines they lay down, for example, is that the word 'risky' is undefined, and is *not* to be used as a synonym for 'dangerous'.

Often the word risk is used simply as a synonym for the *probability* of something unpleasant happening. Although the *likelihood* of some adverse outcome occuring as a consequence of some decision is an important dimension of risk, so also is the *magnitude* of this adverse outcome. When comparing risks it is meaningful to think of risk as a compound – a two dimensional entity comprising both probability and some measure of the magnitude of the adverse outcome. It is this general definition of risk that the Royal Society (1983) put forward commenting that:

> '. . . the Study Group views RISK as the probability that a particular adverse event occurs during a stated period of time, or results from a particular challenge'.

Because the main focus of this paper is that of risks to human life and limb resulting from medical interventions, risks might be viewed as being a range of adverse outcomes (health impairments, reductions in quantity and quality of life) to which can be attached a range of probabilities. Against these probabilistic adverse outcomes will be balanced the probabilistic benefits. By combining probabilities and outcomes (positive and negative) the *expected* outcome can be calculated. By including information on how individuals *value* different health outcomes the *expected value* of a decision strategy can be calculated. Furthermore, by including information on the individuals *attitude to risk* (or risk preference) the *expected utility* of any choice strategy can be computed. All these concepts are familiar territory in clinical decision analysis which forms the background for the discussion of clinical risks and risk taking. (Weinstein *et al*, 1980).

A preliminary task however, is to review the extent to which it is possible to quantify probability and to measure and value health outcomes for use in such analyses.

2 PROBABILITY – 'THERE'S A GOOD CHANCE . . .'

Quantifying the likelihood of events happening is the first stage in assessing risks. Furthermore it is the *presentation* of such probability information that is of key importance to the public's understanding and perception of risk. Is there a consensus on the definition, measurement and comprehension of probability?

As with risk, the definition and interpretation of probability varies between its use in everyday language and the more precise scientific and professional usage. Perhaps the most common synonym for probability is *chance*. The philosopher, A. J. Ayer in an excellent essay on chance reproduced in Dowie and Lefrere (1980), has distinguished between three ways in which chance and probability concepts are used. Firstly, there are statements of the kind that the chance of throwing a double-six with a pair of true dice is one in 36. This involves what is often called a judgement of *prior* probability. If a 'fair' coin is tossed a large number of times the statistical outcome should match the prior odds: a 50/50 chance of heads or tails.

The second use of chance is where there exist no prior odds but probabilities are computed from statistical data from a series of previous events. To state that the chances of any given unborn infant being a boy are slightly greater than 50/50 is to forecast the future from the past.

Finally, there are chance statements which are expressions of what Ayer terms 'judgements of credibility'. One might state that 'there is a good chance that I will not become Prime Minister'. This invokes no prior odds and a probability could not be reliably constructed from previous experience – of mine or anybody elses. It is essentially a judgement which expresses a person's 'degree of belief' that something may occur.

In the context of clinical risks there can be no strict prior probabilities as with the toss of a fair coin. The majority of probability estimates will be projections based on past experience. Thus the ratio of how many times the event *has happened* in the past relative to how many times *it could have happened* could be used as a predictor for its likelihood in the future. Supposing that a series of 1,000 patients had undergone surgery and that 15 died in surgery, then the proportion of patients who had died would provide us with the initial operative fatality probability estimate, this being 15 out of 1,000 or 0.015. Generally the larger the sample size the greater will be the confidence that the observed frequency is a good estimate of the actual probability in the general population. Such probabilities might be termed as being *objective* in as much as they are not judgements of likelihood by a panel of experts but rather a direct appeal to empirical evidence on the frequency of a particular outcome within a known population exposed to that risk.

Although appealing to statistics will usually form the basis of clinical risk estimates, such estimates may be influenced by more subjective elements regarding the degree of belief of an event occuring.

OBJECTIVE OR SUBJECTIVE?

The distinction between objective and subjective probability forms an important divide between two schools of thought on probability. On the one hand the 'frequentist' interpretation is as described above – probabilities are calculated from the frequency of events from the past. However, the subjective or Bayesian (named after Thomas Bayes 1702–1761) school of thought maintains that such reasoning is often applied to events which are not truly repetitive. Holroyd and Collings (1980) explain in the context of establishing betting odds: '. . . one does not find exactly the same horse-race repeated a large number of times, with the same participants and under the same conditions'. Thus for the subjectivist, probabilities are always based on somebody's 'degree of belief' that some event will happen. Everyone has different sets of information from diverse sources and experience which influence their degree of belief and hence their subjective probability estimates.

Although the estimation of objective probabilities from large-scale empirical studies is the mainstay of epidemiological enquiry in this area, decision-making in clinical practice will typically rely on subjective probability estimates formulated by the clinician. For example, the clinical literature may report the mean expected operative fatality rate for coronary artery bypass grafting (CABG) but the surgeon must translate these data into the odds of fatality for his patient of specific age, sex and health state who may not be representative of the population from which the objective estimate was made. The objective assessment would therefore form the baseline 'odds' and the clinician would modify these odds for each individual patient using his judgement based in part on his experience and knowledge of the patient.

A typical illustration of objective versus subjective probabilities in health is smoking. Information on the health risks of smoking is produced by a group of 'experts' and the individual is confronted with an array of 'objective' probabilistic information. Yet a person's subjective estimates of probability may be significantly lower than the objective – this is commonly observed. The smoker may believe that the *specific risks for him or her* are much lower than the reported statistical average because in some way they feel atypical of this statistical population. The degree of belief in the objective probabilities is influenced by past experience; a smoker with healthy parents who have been lifetime smokers is more likely to understate probabilities and therefore perceive lower health risks.

WORDS OR NUMBERS?

Irrespective of whether the estimate of probability is objective or subjective, how best is probability information communicated? A number of surveys have demonstrated that individuals find it difficult to digest numerical constructs such as probability. Prestcott-Clarke and Mostyn (1980) in a survey on public attitudes toward risk found that when confronted with probablistic data '. . . most respondents claimed that such statistics had no useful meaning for them'. In an attempt to 'personalise' such risk information the currency of probability is often verbal; but to what extent is their a consensus or consis-

tency in this numeric-verbal translations?

Lichtenstein and Newman (1967) demonstrated that the interpretation of everyday probability expressions (seems likely, almost certain, etc) is highly ambiguous. When subjects were asked to assign numerical values (between 0 and 100) to 41 different expressions of probability, the range of responses for each word was very large. For example, 'probable' was given values between 1 and 99; 'seldom' between 1 and 47.

Beyth-Marom (1982) recently replicated this experimental approach and the results are presented in Table (1). The 30 probablistic verbal expressions are ordered from 'Not likely' to 'certain'. The range of numerical responses for probabilistic words indicates greatest respondent agreement on the phrase 'certain' and most disagreement and ambiguity was found with the phrase 'can't rule out entirely'. It is interesting to note that such ambiguity has led the US National Weather Service to express forecasts numerically – ie quoting probabilities for rain the next day – rather than verbally. (Murphy and Winkler, 1974). Some interesting parallels might be drawn between those professionals who communicate prognoses on weather and those giving a prognosis on health.

Verbal-numeric ambiguity surrounding probability information is of particular relevance to medicine in the area of informed patient consent to treatment. If doctors have a duty to inform patients about treatment risks then probabilistic information must be communicated. On the one hand patients may not understand numbers, but on the other hand verbal statements may be ambiguous and more misleading than numbers. The paper returns to the problem of treatment risks and informed consent later on.

In summary, it is clear that estimating outcome probabilities for most human activities is an approximate science often laced with a good deal of judgement. 'Objective' probabilities form the basis for forecasting because they are based on past frequencies of particular outcomes. Yet such estimates are often crude and subject to wide margins of sampling error and uncertainty. The clinical decision-maker, faced with the task of placing probabilities on outcomes for patient X can use the available 'objective' evidence and adapt this to the individual case by formulating subjective probability estimates based on experience and knowledge of the patient.

But if the clinician wishes to communicate such probabilities to the patient in order to gain informed consent, the problem still remains of how best to do this – words or numbers? The patient's perception of the risk will be coloured by the way the probability data is presented to him. Weinstein *et al* (1980) in their **'Clinical Decision Analysis'** argue in favour of using probabilities rather than words:

> 'The trouble with semi-quantitive terms is that they can be interpreted differently by different people. We advocate the use of probabilities, not because the numerical assessment in any way adds legitimacy to the opinion of the decision-maker, but because it facilitates communication among decision-makers and permits the decision maker to derive the maximum use from the available information.' (p 39).

This general problem of how best to communicate risk information is discussed further in the section on risk perception below.

Table 1 **Numerical translation (between 0 and 100) of verbal probability expressions: range of expression**

		Range of expression	
		*Interquartile**	
No	*Verbal expression*	*Limits*	*Range*
1	Not likely	5–15	10
2	Very low chance	10–18	8
3	Poor chance	11–25	14
4	Doubtful	16–33	17
5	Low chance	22–34	12
6	Small chance	22–36	14
7	Can't rule out entirely	24–49	25
8	Chances are not great	28–41	13
9	Not inevitable	35–56	21
10	Perhaps	36–53	17
11	One must consider	37–59	22
12	There is a chance	37–60	23
13	May	41–58	17
14	It could be	42–57	15
15	Possible	51–58	7
16	One can expect	51–63	12
17	Reasonable to assume	52–69	17
18	Likely	53–69	16
19	It seems	53–65	12
20	Non-negligible chance	53–67	14
21	It seems to me	54–67	13
22	One should assume	54–68	14
23	Reasonable chance	54–69	15
24	Meaningful chance	63–80	17
25	High chance	75–87	12
26	Close to certain	75–92	17
27	Most likely	78–92	14
28	Nearly certain	83–96	13
29	Very high chance	87–96	9
30	Certain	98–100	2

*Interquartile range is a measure of dispersion and reports the 'middle' 50 per cent of a distribution (ie, between 25 per cent and 75 per cent).

Source: Beyth-Marom (1982).

3 EXPECTED OUTCOME, EXPECTED VALUE AND EXPECTED UTILITY

Expectation is yet another word which has an everyday usage but which also has a precise mathematical usage. Given a statistical distribution of outcomes, the *expected* outcome is that which on average will occur. Before examining how health outcomes can be combined with probabilities for use in clinical decision analysis a brief exposition of mathematical expectation is required.

Suppose there is a Probabilistically Distributed Quantity (PDQ) such as the number of days a person will remain 'ill' following a minor operation. From previous experience it is estimated that there is a 20 per cent chance of getting well after 1 day, 25 percent after 2 days, 40 per cent after 3 days and 15 per cent after 4 days. The expected outcome is a weighted average of the PDQ (1, 2, 3, 4 days) for which the corresponding weights are the probabilities (0.2, 0.25, 0.5, 0.15). Thus the expected outcome (E) is:

$$E = (0.2)1 + (0.25)2 + (0.4)3 + (0.15)4$$
$$= 2.5 \text{ days}$$

In general it is possible to compute the expected (mean) outcome for any PDQ as being the sum of the products of probabilities (P_i) and outcomes (X_i) such that expected outcome is:

$$E(x) = P_1X_1 + P_2X_2 + \ldots P_nX_n = \sum_{i=1}^{n} P_iX_i$$

A common application of such expectancy calculations in medicine and elsewhere is with life-expectancy where the probabilities of survival to particular years following some event are used to calculate expected (mean) survival. Indeed, length of life or the probability of death serves as the main currency in calculations of human risks. This is largely due to the fact that mortality and its causes remains one of the 'hardest' pieces of epidemiological data.

But to present risk data exclusively in the form of mortality probabilities or life expectancy is to imply that the avoidance of death is the only source of concern. Yet if the focus of interest is *health risks* then the measure of outcome being combined with probabilities must embrace *qualitative* as well as *quantitative* aspects of life. There is growing recognition that the measurement of quality of life is important when quantifying treatment outcomes; as Evans *et al.* (1984) state '. . . we have come to expect more of treatment than mere survival'. Life and death are the black and white extremes while many of the risk decisions we face encompass the interim shades of grey – degrees of morbidity which reduce the *quality* of life.

Calculating expected outcomes in medicine involves a process of attaching probabilities to a range of health outcomes. Such outcomes might be in terms of simple symptom-based statement on physical and mental functioning, relating such measures to 'benchmarks' of normality. For evaluative purposes however, the question then becomes one of determining how individuals *value* different health states relative to each other. In other words, attempting to place some common currency or scale on health outcomes so that they can be systematically compared in a decision context.

FROM OUTCOME TO VALUE

To illustrate the principle of moving from expected health outcomes to expected values consider the following example taken from Weinstein *et al* (1980). A patient has vascular insufficiency (gangrene). A decision must be taken about whether to operate immediately – amputating the leg below the knee. There is a 99 per cent patient survival rate assigned to the operation. An alternative strategy is to wait. If they wait there is a 70 per cent chance that the leg will heal naturally and there will be a total cure. However there is a 30 per cent chance that the gangrene will spread and that later on the surgeon will have to amputate *above* the knee with only a 90 per cent survival rate. What would you choose?

The problem is set out in the form of a decision tree in Figure 1(i). The four basic outcomes A, B, C and D are the end-points of the two strategies. In order to evaluate each strategy and hence decide which is preferable some relative values must be assigned to the outcomes. One might first of all *rank* the outcomes in order of death, above-knee amputation, below-knee amputation and total cure. But *by how much more* is, say, below-knee amputation preferred to above-knee amputation? Because the outcomes are not all in the same units (eg life-expectancy) it is necessary to convert them into some value scale to enable the calculation of expected values for each strategy.

In Figure 1(ii) illustrative values have been attached to the four outcomes. If a value of 1 is attached to 'total cure' and a value of 0 to death the problem becomes one of assessing the quality-of-life of above-knee and below-knee amputation relative to the two extreme values which form the boundaries of this value scale. Begging the question, for the moment, of *how* one gets individuals to make such judgements, but assume they attach a value of 0.9 to below-knee amputation and 0.8 to above-knee amputation. It now becomes possible to calculate the expected value for each strategy – wait or operate now. The conclusion from the calculations in Figure 1(ii) is that the expected value is higher (0.92) if the individual waits rather than accepts below-knee amputation now (0.89). (You might like to check whether this is the strategy you would have chosen and compare the pay-off when the values attached to the outcomes are varied.)

This simple example of clinical decision analysis demonstrates the way in which, by incorporating individuals' valuations to health outcomes the choice criterion is changed from *expected outcome* to *expected value*.

FROM VALUE TO UTILITY

An important ingredient to the analysis of risk in decision making is that of the individuals' *attitude to risk*. Some individuals may have a positive preference for taking

risks or gambling while others will prefer to avoid or minimise risk taking. Therefore value or utility is assigned to the *process* of decision making when it involves risks, in addition to the *outcome* utility or value of the consequence of a decision.

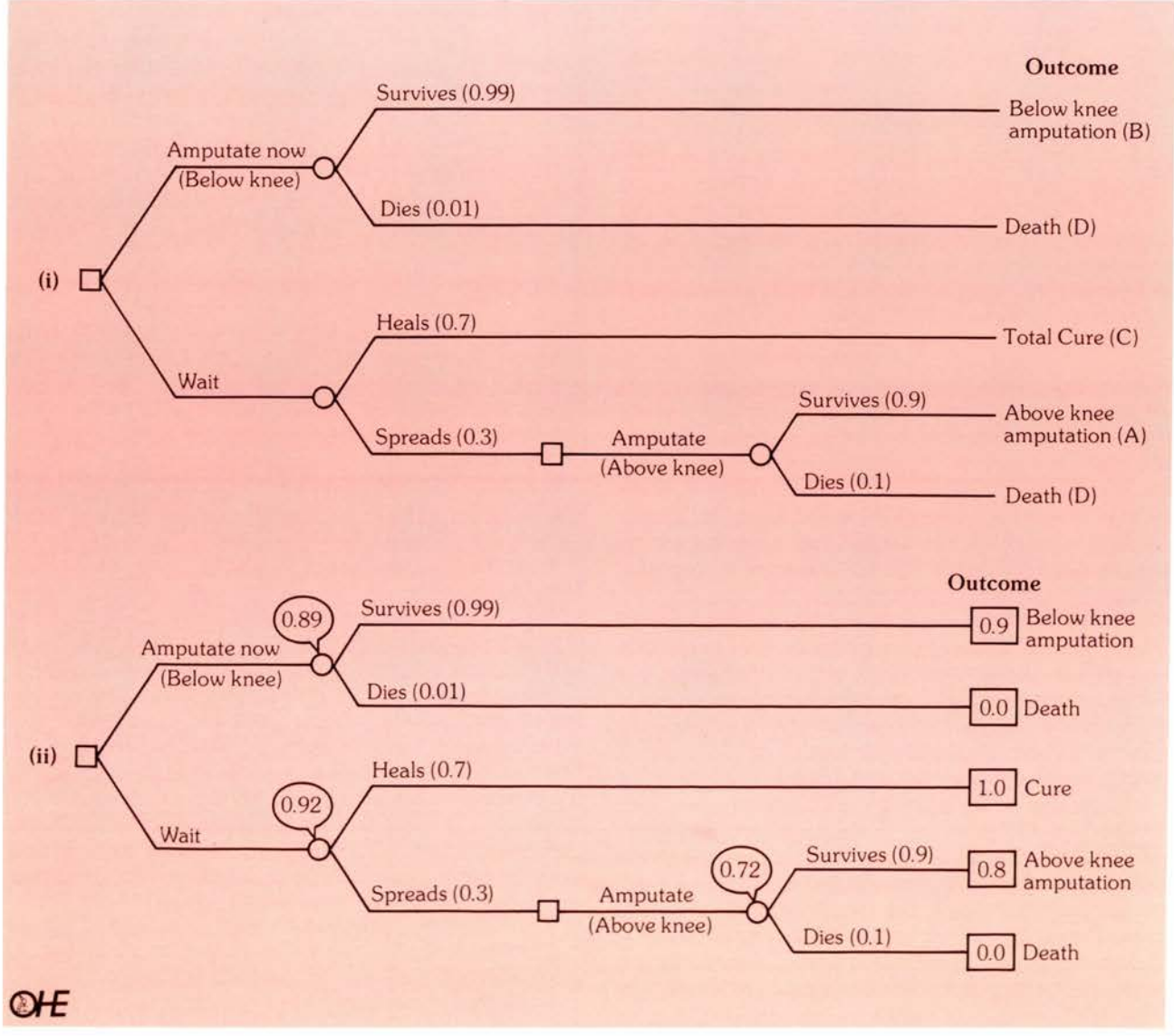
Faced with the problem of evaluating a choice problem (eg between treatments) in terms of expected utility, attitudes to risk must be incorporated. This stems from the pioneering work on expected utility theory by von Neumann and Morgenstern (1947). Attitudes can be divided into three categories: (1) *Risk averters*, who prefer to avoid risks; (2) those who are *risk neutral* and indifferent about risk taking and (3) *risk lovers* who have a preference for risk taking.

The risk averter, for example, is a person who would prefer to buy insurance at a fixed (certain) premium rather than run the risk of financial loss due to fire or theft. The expected outcome financially of insuring or not over the years *may* be the same (he pays as much in premiums over the years as he might have lost if the risk had materialised), but utility is derived from the *process* of not taking the chance. The extent to which the individual values this certainty is reflected in how much he is willing to pay for this insurance.

Attitudes to risk are preferences which both doctors and patients will have when faced with treatment decisions which have differing probabilities and pay-offs. Consider a treatment decision problem where the only outcome of interest is the length of survival. The treatment is a gamble offering a 50/50 chance of survival for 50 years or immediate death. The expected outcome of the gamble is therefore 25 years [0.5(0) + (0.5(50)]. Suppose that the alternative of no treatment offers a certainty of 15 years. If the outcomes are identical except for longevity, which strategy should be chosen; treatment or not?

Again the answer depends upon the decision-maker's attitude to risk. Although the expected outcome of gambling on treatment is higher, the expected utility is the decision criterion of interest because this includes pre-

Figure 1 **Clinical decision analysis of vascular insufficiency (Gangrene) (from Weinstein *et al*, 1980, p 187)**



1 In diagram (i) the expected outcomes for the two strategies involve comparing health outcomes other than life expectancy. In diagram (ii) values have been assigned to outcomes and hence the expected value of the two strategies can be determined. (Note, however, that this does not include information on attitude to risk which will form an input to expected utility).

2 Decision nodes are □ and chance nodes are ○

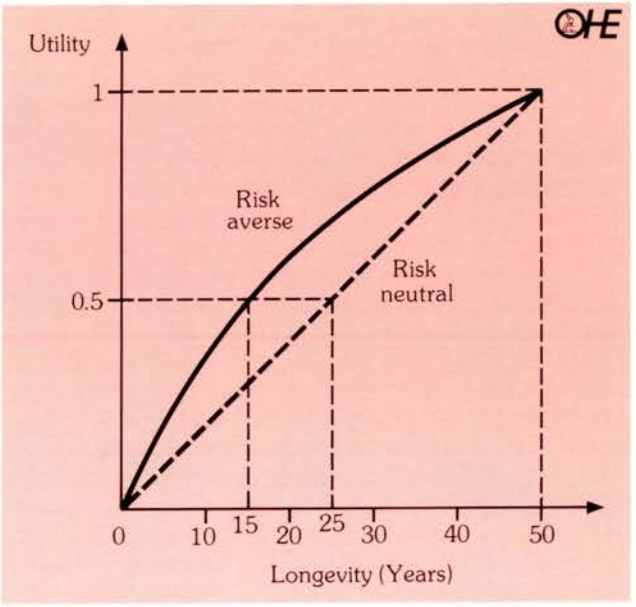
ferences regarding the *process* of gambling or risk taking. Depending upon how much the individual dislikes the idea of the gamble the more likely he is to accept the certainty, even though it has a lower expected pay-off. The individual is said to be *risk averse*. The degree of risk aversion with respect to longevity is a matter of personal preferences. As Weinstein *et al* note:

> 'It may depend on one's age, family responsibilities, and many other factors. For example, an individual with young children may be very risk averse in order to ensure sufficient time to provide for their future, while a young single person may be less risk averse. In addition, people may assign greater value to more proximate years of life than to the more distant years; this preference would appear as a source of risk aversion.'

Figure (2) illustrates the problem and compares two individuals who differ in their attitudes to risk. Someone who is risk neutral does not derive disutility from the treatment lottery and expected utility can be equated with expected longevity. However, the risk averter demonstrates a preference to avoid risk taking such that the 'certainty equivalent' of the treatment gamble *to him* is 15 years. In other words, if the no-treatment option offered 16 years of life with certainty he would prefer this to the treatment gamble even though the life-expectancy of the treatment gamble is higher.

Although treatment choices are likely to be less dramatic than this example, it serves to illustrate the principle that many individuals are not indifferent regarding their preferences on risk taking. It follows from this that if doctors or surgeons, when faced with such treatment decisions, choose that which maximises life expectancy, this may not be the strategy which maximises the patient's expected utility. They may not wish to gamble on the lottery even if the expected pay-off in life years is higher than not gambling. This question is further examined in the section on risk balancing below.

Figure 2 **Measuring expected utility on the longevity scale: risk averse and risk neutral individuals**



The treatment is a 50/50 gamble of 50 years or 0 years; the expected pay-off being 25 years. ((0.5)50 + (0.5)0). For the risk neutral person the expected outcome is the same as expected utility: he is indifferent about the process of gambling in this way. For the risk averter the certainty equivalent of the treatment gamble is 15 years: he would be indifferent between no treatment if it offered (in this case) a certainty of 15 years and the expected treatment pay-off of 25 years.

Source: Weinstein *et al* (1980)

4 PROBABILITY AND THE VALUE OF HEALTH

In the gangrene example discussed above, the decision strategy hinged on whether values could be attached to non-fatal impairment outcomes. Attempts to measure and value health status is a growing area of research gaining wide application in economic evaluations of health care activities. A detailed review of the various approaches and accounts of many of the available instruments can be found in Culyer (1982) and more recently in Teeling Smith (1985) and Torrance (1986). Particular effort has gone into researching ways in which an overall health index might be constructed which could locate a specific health state on a continuum between 0 (= death) and 1 (= full health). Rosser and Watts (1972), for example, have developed a method where all health states between death and full health can be described in terms of the two dimensions of distress and disability. Having elicited relative values for various combinations of these two parameters they were able to devise an index for rating health states between 0 and 1.

A number of researchers have attempted to build similar indices and profiles for measuring health. Yet even if the range of health states can be described or quantified one of the basic problems is how to determine the *value* of one health state relative to another as perceived by the individual, and more generally by the public at large.

A method of valuing states which is more consistent than others with the expected utility framework, is the **standard gamble** technique. Unlike scaling methods where health states are explicitly valued, the standard gamble method is an implicit valuation technique because it is based on an individual's responses to a hypothetical decision situation; the choice of strategy implying the value of a health state to that individual.

THE STANDARD GAMBLE APPROACH

Consider an individual who is faced with the choice of remaining in a particular (less than full) health state, versus a gamble on treatment which could fully restore health or could result in death. The choice is mapped out in Figure 3. Suppose in state (A) the person has severe angina, he could live on with the pain or gamble with bypass surgery which could remove the angina and fully restore his health (state H) which occurs with probability (P) or it could prove fatal (state D) with a probability of (1–P).

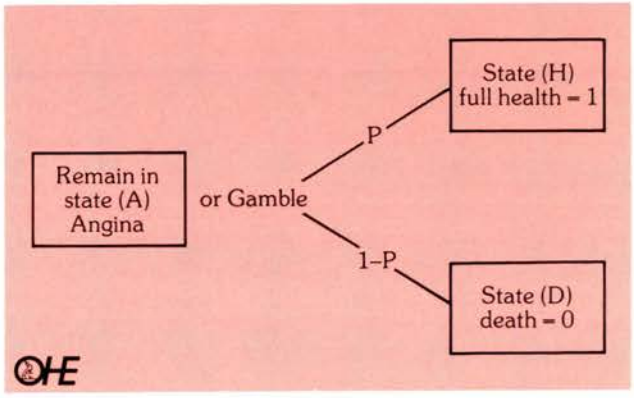
If it is assumed the utility value of full health is 1 and of death is zero the object of the exercise is to locate state (A) with angina on this utility scale. If the probability of restoring full health (P) is varied, then there will come a point where the individual is indifferent between living on in state A and taking the gamble of surgery. In other words the expected utility of state A would be equal to the expected utility of the gamble:

$$U(A) = pU(H) + 1\text{–}pU(D)$$

$$U\,(\text{certainty}) = U\,(\text{gamble})$$

Where U(H) and U(D) are the utilities associated with full health and death. When the 'best' outcome of the gamble is full health, U(H) is arbitrarily set at 1.0 and when death is the 'worst' outcome, U(D) is set at 0; and thus U(A) = p. In other words the probability at which the individual is indifferent between the gamble (surgery) and the certainty (state A) is the utility measure of that certainty (state A).

Figure 3 **A standard gamble choice: angina and coronary bypass surgery**



The probability (P) on the gamble is varied until the point is found where the individual is indifferent between the certainty and the gamble. The expected utility from the gamble being the same as the certainty. This level of (P) is then equivalent to the individuals utility rating of state (A), the process having determined where state (A) lies between the values of 1 (state H) and 0 (state D).

Example: If the probability of indifference is p = 0.9 then the utility assigned to state (A) would be 0.9. The more he wants to move from the discomfort of state (A) the more he is prepared to gamble; hence the lower is (p) and the lower the utility rating of state (A).

If the individual perceives the angina as particularly undesirable his indifference will decrease – he will be prepared to accept a greater probability of mortality to escape state (A) – and hence the utility rating of state (A) will be lower.

Another technique for eliciting utility values for health states very similar to that of the standard gamble is the technique of **time trade-off** developed by Torrance and colleagues at McMaster University in Canada (Torrance, 1986). With this method, rather than varying the probability associated with a treatment gamble, the technique is to vary the length of time in each health state with the choice of treatment strategy again revealing the implied utility value of any given health state.

The similarity between the 'standard gamble' and the 'time trade-off' approach is discussed in Gafni and Torrance (1984). This article is an excellent exposition of the concepts of time preference and risk preference as they apply to health gains and losses. In particular the authors argue that an individuals' risk attitude in health is the result of three effects – a quantity effect, a gambling effect and a time preference effect. Therefore **risk aversion** can be explained in terms of:

(i) *quantity effect:* diminishing marginal utility with respect to, for example, life years. A conventional economic principle which translates broadly into the maxim

'the more you have of anything the less you want any more of if'.

(ii) *gambling effect:* basically the 'process utility' associated with risk taking is negative for a risk averter.

(iii) *time preference effect:* generally people wish to have beneficial things sooner rather than later (positive time preference rate). Therefore life-years occuring further into the future would be valued less than those occuring earlier.

All these elements would therefore combine to explain the concavity of the utility function in Figure 2, and are concepts intrinsically bound-up with the way in which we value health gains and losses in situations of choice involving risk.

QUALITY-ADJUSTING LIFE EXPECTANCY

Outcomes from treatments and other health influencing activities have two basic components: quantity and quality of life. Life expectancy is a traditional measure with few problems of comparison. Furthermore the previous section illustrated the ways in which values can be assigned to health states between 'full health' and 'death' using the 'standard gamble', the 'time trade-off' or other techniques for scaling health states. Weinstein and Stason (1980) demonstrate how these two components of quantity and quality could be aggregated into a single measure – the Quality Adjusted Life Year or 'QALY' as it has become known.

The idea of the QALY is very simple. Because not all years experienced will be at 'full health' these can be adjusted downwards to standardise for quality. A year at full health could count as 1.0 QALY, but two years, each valued at 0.4 would count together as 0.8 QALY. In principle then, the consistent application of such QALY measurement with clinical outcomes would allow treatment strategies with very different quantity and quality components to be compared using a common unit of account. A number of authors have advocated the calculation of QALYs such that resource allocation decisions might be guided by the relationship between inputs (costs) and outcomes (QALYs). [For cost-per QALY calculations see Torrance (1986) and Williams (1985)].

From the perspective of clinical decision analysis the expected outcomes of a given choice strategy can be translated into a probabilistic value measure – expected QALYs. It provides a useful framework for analysing treatment 'gambles' where, to achieve some gain in quality of life (eg reduced pain and discomfort), the individual accepts the small probability of fatality associated with the treatment in order to gain the improvement in quality of life. Obvious examples of this include elective surgery such as hernias, haemorrhoids and hip replacements.

More generally the expected gains and losses in QALYs could be applied to any area of human activity where risks to life and limb are involved. Provided values can be attached to the relevant range of health states, outcomes associated with different choice strategies could be expressed and compared in terms of expected QALYs.

In the following section it is shown that the practice of risk assessment generally falls well short of the principles and concepts discussed in this section. Estimates of risks, both in medicine and more widely, are focussed on fatality probabilities, This being so, little emphasis is placed on the task of valuing risks to *health* which are non-fatal but provide an input in many decision contexts.

5 RISK ASSESSMENT: A HAZARDOUS BUSINESS

The assessment of risks is a complex issue which has recently been the subject of a study group report from the Royal Society (1983). The aim of this section is briefly to review some of the issues that arise in attempts to assess risks, discussing both the availability and usefulness of event-frequency data and in particular the problem of finding an appropriate measure of exposure to act as a denominator in risk calculations. The various ways in which risk data are calculated and presented are compared and contrasted, with a review of some recent attempts to construct a simple logarithmic risk scale for presenting such data.

Methods for assessing risks from a number of activities, including medicine, are discussed. The purpose being to identify general problems in risk assessment and consider these in the specific context of clinical risks.

SOME DEFINITIONS AND TERMINOLOGY

The excellent report on risk assessment produced by the Royal Society (1983) was an attempt to pull together the many aspects of risk – its assessment, perception and management – and to explore the way in which risk terminology and concepts are applied in different professional disciplines. In the opinion of the study group the consensus use of risk terminology was as follows:

> 'The general term used to describe the study of decisions subject to uncertain consequences is RISK-ASSESSMENT. It is conveniently subdivided into RISK ESTIMATION and RISK EVALUATION. The former includes:
>
> (a) the identification of the outcomes;
> (b) the estimation of the magnitude of the associated consequences of these outcomes; and
> (c) the estimation of the probabilities of these outcomes.
>
> RISK EVALUATION is the complex process of determining the significance or value of the identified hazards and estimated risks to those concerned with or affected by the decision. It therefore includes the study of risk perception and the trade-off between perceived risks and perceived benefits. RISK MANAGEMENT is the making of decisions concerning risks and their subsequent implementation, and flows from risk-estimation and risk-evaluation.'

In keeping with this taxonomy the main aim of this section is to examine the various ways in which risk outcomes have been identified and presented, paying particular attention to the problems of estimating probabilities from frequency data.

Cohen and Pritchard (1980) argued that risks can be classified into three basic categories:

(a) Risks for which statistics of identified casualties are available.
(b) Risks for which there may be some evidence, but where the connection between suspected cause and injury to any one individual cannot be traced (eg cancer long after exposure to radiation or a chemical).
(c) Experts' best estimates of probabilities of events that have not yet happened.

Estimating the likelihood and consequences of each type of risk obviously raises quite different problems. Estimating risks from group (a) is an exercise in predicting the future from the past. A frequency distribution for an adverse event such as road traffic mortality by type of vehicle and road can be combined with data on the volume and type of journeys in a given period of time to produce a measure of risk; in this case the probability of fatality per unit of time.

This type of reasoning and risk estimation can be applied to medical and surgical activities where details of 'casualties' have been recorded for a known population exposure. Following on from the discussion of probability, such estimates might be termed 'objective' because they are empirically based forecasts rather than subjective judgement by either a lay-person or an 'expert'.

The problem with estimating risks from group (b) is that of 'cause and effect'. This is particularly a problem when there is a substantial time lag between the individuals exposure to the risk and the subsequent materialisation of the adverse consequences. Examples include exposure to radiation and other carcinogenic risks which may take many years to result in cancers. Despite the rigour of epidemiological enquiry, the strongest statistical bond that ever exists is one of *association* rather than causation. With this group of risks then, the measure chosen to present risk data must allow for the time lag between suspected cause and effect so that the relationship can be further investigated. One approach, for example, is to present risk data in the form of reductions in life-expectancy.

The problems of estimating risks from group (c) are essentially those of asking: how probable is the possible? To use a topical example, supposing that prior to the Chernobyl disaster in the Soviet Union you had been given the job of estimating the likelihood of such a nuclear power plant disaster. It had never happened before so how probable was it? One technique used to estimate such probabilities is called **fault free analysis**. An event like a nuclear core melt down can be broken up into a number of contingent events, and the probability of each chain in the link occuring might be estimable from previous experience. If one can calculate the probability of each domino in the chain falling then it becomes possible to calculate the probability of the end-point being achieved. Fault free analysis is discussed more fully in Fischoff *et al.* (1981).

For the most part, medical and surgical risks fall into group (a) where some data on outcomes from a specific number of interventions are available. Two particular risk measures to be examined in more detail are surgical case fatality rates and adverse reaction rates from pharmaceuticals.

ESTIMATION AND PRESENTATION

Given that the purpose of risk estimation is to provide the decision maker with risk data in a comprehensible for-

Table 2 **Annual risk by selected road user, UK 1983**

	Billion vehicle Kms	*Billion passenger Kms*	*Number of fatal casualties*	*Fatalities per billion*	
				vehicle Kms	*passenger Kms*
Motor cycles	6.2	7.0	963	155	138
Pedal cycles	5.0	5.0	323	65	65
Cars and taxis	220.0	407.0	2,019	9	5
Buses and coaches	3.3	42.0	38	12	1

Source: Department of Transport.

mat, the question of how best to estimate cannot be divorced from that of how best to present risk information. Given the diversity of professional applications of risk assessment it seems unlikely that a universal 'gold standard' for the way risk data should be presented will emerge. Regarding general guidelines however, Lord Rothschild (1978) voiced a strong opinion on the way risk data should be presented:

> 1 'Is the risk stated in a straightforward language that I can understand such as one in 1,000? If not, why not?'
> 2 'Is the risk stated per year, per month, per day or per some period of time? If not, I shall ignore the information.'

and on these points Inman (1984) concluded that: 'Few public statements by government agencies, politicians, pressure groups or journalists satisfy these two simple requirements.'

There are basically three methods for presenting risk information:

(i) Adverse outcomes (eg fatalities) per number of events per time period.
(ii) Average reduced life-expectancy (if possible, quality adjusted) resulting from some given period of exposure to a particular hazard.
(iii) Risk equivalents: catalogue of events or amount of activities all of which increase probability of specific adverse outcome by same amount in given time period.

An illustration of the first type of measure is road traffic accident and mortality statistics by type of vehicle and other characteristics which can be combined with a measure of exposure – either vehicle kms travelled or passenger kms – to produce a measure of risk (eg annual fatalities per billion vehicle kms travelled). Such data are presented in Table 2.

A number of points can be made about the choice of exposure measure. If the total number of fatalities were simply taken as the risk indicator without allowing for relative exposure then it would be tempting to conclude that 'cars and taxis' are the most risky mode of travel – having an annual fatality rate of 2,019. However, different risk relativities soon became apparent when exposure is examined. Car and taxi *passengers* had approximately 58 times greater exposure (passenger kms) than motor cyclists and yet only incurred just over twice the number of fatalities. Note however, that the *choice* of exposure measure in Table 2 – vehicle kms or passenger kms – will be important where there is differential vehicle *occupancy*.

Further examples of this type of risk presentation are commonly found in analyses of fatal accidents at work and during sporting activities. Table 3 lists a variety of sporting activities with estimates of their attendent fatality risks. For comparative purposes the most useful exposure measure is some indication of time spent 'doing the sport' to produce risk measures of deaths per participant-year or participant-hour. However, accurate data on activity and time exposure is scarce – often limited to club or association membership figures – and therefore measures of exposure will generally have wide margins of error associated with them. It is difficult, for example, to compare activities when risks are calculated in different time units; how many participant hours are there in a participant year? Moreover, the choice of time unit may actually change the ranking of risks when they are compared, as the Royal Society (1983) note:

> '. . . if risk A is the risk of being killed by fire per unit of time of staying at home and risk B is that of being killed by a fall during rock climbing, then if we compute deaths per person hour, rock climbing is riskier than fire; but if we compute deaths per person year, the order is reversed.'

Choice of exposure measure is no less a problem when computing clinical risks. Consider surgical fatality risks; is it more meaningful to calculate fatalities per operation, in a given time period, or fatalities per operating-hour? It really depends on what question is being asked and for what purposes the information is needed. In general it does indicate that not all risk estimates are as robust and 'objective' as they appear at first glance. The analyst's

Table 3 **Risk of death in sporting activities**

		*Deaths per 10^6 participant years**
Cave exploration	(US, 1970–78)	45
Scuba diving	(UK, 1970–80)	220
	(US, 1970–78)	420
Glider flying	(US, 1970–80)	400
Power boat racing	(US, 1970–80)	800
Hang-gliding	(UK, 1977–79)	1,500
	(US, 1978)	400 to 1,300
Sport parachuting	(US, 1978)	1,900
		*Deaths per 10^6 participant hours***
Amateur boxing	(UK, 1946–62)	0.5
Skiing	(US, 1967–68)	0.7
	(France, 1974–76)	1.3
Canoeing	(UK, 1960–62)	10.0
Mountaineering	(US, 1951–60)	27.0
Rock climbing	(UK, 1961)	40.0

* Based on numbers of participants and deaths per calendar year, without allowance for hours actually spent in the activity.
** Based on approximate estimates of participants' hours per year spent in the activity.
Source: Royal Society (1983)

choice of exposure measure may markedly change the relativities of risks being compared.

The second method of presenting mortality risk data, which has been used fairly widely in work hazard comparisons is to calculate reduced life-expectancy as the result of some period of exposure to a hazard. Table 4 is taken from Reissland and Harris (1979) and compares age-specific average life-expectancy lost due to one years exposure to hazards in a number of industries. Thus for someone aged 20, a year deep sea fishing would reduce his/her average life expectancy by 51.4 days. The equivalent reduction for a year's work in the nuclear industry (at the maximum allowed dosage rate of 5msv) would only be 4.6 days.

Reissland and Harris (1979) argue that the reduced life-expectancy method is more meaningful than presenting fatal accident rates per person-year. In particular they note that '. . . it is a useful measure of risk because it permits a comparison of "instant" accidental death with the delayed effects of radiation'. The emphasis is not simply on lives lost but on the number of life-years lost. Such an outcome measure seems more in keeping with the health care outcome measures discussed above. In medicine nobody can 'save' a life, only extend it. Losses and gains in life-years being quantified relative to normal life expectancy which will be dependent upon age.

This approach to risk estimation also lends itself more easily to the incorporation of morbidity events – qualitative aspects of life and the idea of expected QALYs discussed in the previous section. Examining occupational risks, the International Commission on Radiological Protection (1977) attempted to construct an **Index of Harm**. They proposed to compare occupations on the basis of the total lost years of life or normal activity (man-years per thousand man-years of exposure) from deaths, non-fatal accidents, and diseases of occupational origin – combining these into an index of harm.

The index included morbidity events (if attributable to the job) which reduced 'normal activity'. An arbitrary weighting factor of 10 was applied to the loss of a year due to death relative to the loss of a year due to other factors. Therefore, although the principle of including non-fatal health effects into a general risk index is encouraging, the approach to determining the relative value of such events seems crude. On this general subject of values in risk assessment the Royal Society (1983) commented:

> 'At present, the relatively simple system of describing and estimating each type of risk separately seems preferable, leaving the judgement of how to weight each detriment (which must, of course, be reached by appropriate and informed open discussion) to those responsible for determining practical programmes.'

This general advice, although it offers no guidelines about how relative values might be attached, does underline the importance of including all detrimental effects with an appropriate weight into measures of risk.

The third method of presenting risk information is to produce a compendium of given quantities of various activities, all of which are estimated to be equivalent in terms of fatality probability. Table 5 is from Wilson (1979) and lists a number of activities all of which are estimated to increase the probability of fatality in a year by 1 in a million. Thus smoking 1.4 cigarettes is estimated to be as risky as living within 5 miles of a nuclear reactor for 50 years which is equivalent to travelling 10 miles by bicycle

Table 4 **Days of life expectancy lost as a result of hazards in the nuclear industry compared with hazards in other industries**

	Age in years at beginning of exposure				
	20	*30*	*40*	*50*	*60*
One year at risk in:					
Deep-sea fishing	51.4	41.6	31.9	22.8	14.9
Coal mining	5.7	4.6	3.6	2.5	1.7
Coal and petroleum products	4.1	3.3	2.6	1.8	1.2
Railway employment	3.5	2.9	2.2	1.6	1.0
Construction	3.5	2.8	2.1	1.5	1.0
All manufacturing	0.7	0.6	0.5	0.3	0.2
Paper, printing and publishing	0.5	0.4	0.3	0.2	0.1
Radiation work at 50 mSv/year	4.6	2.7	1.3	0.5	0.1
Radiation work at 5 mSv/year	0.4	0.3	0.1	0.1	0

Note: mSv = millisievert is a measure of radiation dosage.
Source: Reissland and Harris (1979)

Table 5 **Risks estimated to increase chance of death in any year by 0.000001 (1 part in a million) (USA)**

Activity	*Cause of death*
Smoking 1.4 cigarettes	Cancer, heart disease
Drinking 0.5 litres of wine	Cirrhosis of the liver
Spending 1 hour in a coal mine	Black lung disease
Spending 3 hours in a coal mine	Accident
Living 2 days in Boston or New York	Air pollution
Travelling 6 minutes by canoe	Accident
Travelling 10 miles by bicycle	Accident
Travelling 150 miles by car	Accident
Flying 1,000 miles by jet	Accident
Flying 6,000 miles by jet	Cancer caused by cosmic radiation
Living 2 months in Denver on vacation from New York	Cancer caused by cosmic radiation
Living 2 months in average stone or brick building	Cancer due to natural radioactivity
One chest X-ray taken in a good hospital	Cancer caused by radiation
Living 2 months with a cigarette smoker	Cancer, heart disease
Eating 40 tablespoons of peanut butter	Cancer caused by Aflatoxin B
Drinking Miami drinking water for 1 year	Cancer caused by chloroform
Drinking 30 12oz cans of diet soda	Cancer caused by sacharin
Living 5 years at site boundary of a typical nuclear power plant	Cancer caused by radiation
Drinking 1,000 24oz soft drinks from recently banned plastic bottles	Cancer from acrylonitrile monomer
Living 20 years near PVC plant	Cancer caused by vinyl chloride
Living 150 years within 20 miles of a nuclear power plant	Cancer caused by radiation
Eating 100 charcoal-broiled steaks	Cancer from benzopyrene
Risk of accident by living within 5 miles of nuclear reactor for 50 years	Cancer caused by radiation

Source: Wilson (1979)

Table 6 **Logarithmic scale of risk levels**

Risk level	*Range*		
1	1 in 1	–	1 in 9
2	1 in 10	–	1 in 99
3	1 in 100	–	1 in 999
4	1 in 1,000	–	1 in 9,999
5	1 in 10,000	–	1 in 99,999
6	1 in 100,000	–	1 in 999,999
7	1 in 1,000,000	–	1 in 9,999,999
8	1 in 10,000,000	–	1 in 99,999,999

Source: Inman (1984)

and so on.

Such compendiums of comparable risks, if reliably estimated, can provide a useful basis for establishing the relativity of risks, enabling individuals to place risks from new technologies and developments in the context of everyday risks which are commonly run. During the Windscale Inquiry the Hon Mr Justice Parker regarded this type of approach as being important when explaining risks to the lay person:

> 'I have no doubt that the best way to explain the degree of risk to the public is to give a broad range of comparables', (Parker, 1978).

In the same year the value of comparison was also emphasised by Lord Rothschild in his Reith Lecture:

> 'There is no point getting into a panic about the risks of life until you have compared the risks which worry you with those that don't, but perhaps should.' (Rothschild, 1978).

RISK LEVELS AS AN ALTERNATIVE PRESENTATION

Perhaps an appropriate conclusion to draw is that estimating risks is a hazardous business. The reliability of event-frequency and outcome data is variable and the latter is often limited to mortality. Choice of exposure measures, where data permits, introduces further uncertainties about probability estimates for use in risk comparisons. In addition to this the public will generally have difficulty in comprehending the numerical presentation of very small probabilities.

Urquhart and Heilman (1984) have argued that, at best, our estimates of risks should be viewed as broad *orders of magnitude* and they proposed a simple risk scale based on a logarithmic translation of risk data. The simple risk level scale is illustrated in Table 6. Thus a probability of 1 in 1 (certainty) to 1 in 9 would be risk level 1, while 1 in 10 to 1 in 99 would be risk level 2 and so on. Thus the risk level is the logarithm (base 10) of the upper limit of the risk denominator and each risk level goes up by a factor of 10. The effect of the logarithmic scale is therefore to condense the very wide range of risks into a simple 8 point scale covering the odds of 1 in 1, to 1 in 10 million.

Similar to Inman's (1984) use of the Urquhart-Heilman risk scale, the mortality risks by cause of death for the population of England and Wales in 1984 have been calculated to demonstrate the scale and these data are presented as risk levels in Table 7. During that year about one in a hundred people died and therefore deaths from any cause fall into risk level 2. Grouping the data by risk level gives the reader the order of magnitude of risk, cancers and heart disease at level 3, while whooping-cough and measles at level 8.

The Urquhart-Heilman scale is a simple means of presenting complex risk data and a useful technique for grouping together risks of the same order of magnitude. Its main value would appear to be for comparing different risks which have been calculated on the same frequency-exposure data. Yet regarding its understandability it remains unclear whether the public would comprehend risk level 8 anymore than they would understand probabilities expressed as 1 in 10 million. Replacing numbers with their logarithms may simplify presentation but it cannot be guaranteed to improve the lay-person's comprehension of probability and risk data.

Table 7 **Selected mortality risk levels; England and Wales 1984**

	Number of deaths in 1984	*Probability of mortality (Deaths/Pop)*[1]	*Urquhart-Heilman risk level*
(All causes)	566,881	1.0×10^{-2}	2
Cancers	140,101	2.8×10^{-3}	3
Coronary heart disease	157,506	3.2×10^{-3}	3
Strokes	14,211	2.9×10^{-4}	4
Diabetes	6,369	1.3×10^{-4}	4
Asthma	1,764	3.5×10^{-5}	5
Cirrhosis	2,280	4.6×10^{-5}	5
Ulcers (stomach and duodenum)	4,483	9.0×10^{-5}	5
Pregnancy*	52	1.4×10^{-6}	6
Acute rheumatic fever	2	7.0×10^{-6}	6
Influenza	346	2.9×10^{-6}	6
Syphilis	68	2.0×10^{-7}	7
Measles	10	2.0×10^{-8}	8
Whooping-cough	1	4.0×10^{-8}	8

* Calculated on female population only

1 Population denominator = 1984, mid-year for England and Wales. Thus for 100 people chosen at random from 1984 population, one of them will die from some cause during the year.

6 ASSESSING MEDICAL AND SURGICAL RISKS

Diseases, desperate grown,
By desperate appliance are relieved
Or not at all

Shakespeare, *Hamlet*

This section examines the ways in which the risks of medical and surgical treatments are assessed. In particular it focuses on the risks of prescription medicines and surgery. The review of pharmaceutical risks is prefaced by a brief discussion of the safety regulations surrounding the industry and the introduction of new products onto the market. The review of surgical risks – mainly measured in case fatality rates – is placed in the context of the probabilistic nature of diagnosis and decisions concerning whether to operate or not.

A theme that is central to the discussion of clinical risks is that of *risk balancing* as an evaluative device in treatment decisions. The levels of treatment risk that are 'accepted' are in proportion to what is at stake – balancing the risks of action against those of inaction. While the introductory sections discussed the framework for examining the overall expected outcome from treatment strategies – summing the positive and negative health effects with their attendant probabilities – this section deliberately focuses on the *negative aspects*, the chance of treatments producing adverse effects in varying degrees.

POCHIN'S COMPENDIUM

Sir Edward Pochin has been a regular contributor to the literature on the risks associated with medical and surgical interventions. (Pochin 1975, 1981, 1982 for example). He has reviewed a number of routine data sources and specific studies to produce compendia of fatality risks – a compilation of which is presented here as Table 8. Pochin's chosen mode of calculation and presentation is that of estimating the relative frequency of fatality associated with each procedure. Thus prescription medicines are estimated to carry a fatality probability of one per million scripts (1×10^{-6}). This compares, for example, with the risks of child-bearing being one death in every 10,000 maternities.

It is difficult to make any comment about the risks of one procedure *relative* to another because there is no common measure of exposure (such as time) and because no indication is given regarding the time-lag between procedure and fatality. Again, the question about which exposure measure is more relevant is dependant upon the question being posed. To repeat an earlier question, should one compare surgical risk in terms of fatalities per operation or per operating-hour? Logically, the longer the patient is in theatre and anaesthetised, the greater the exposure to the possibility of something going wrong. Then again, it might be argued that risks such as these are 'fixed' and it is the event (the operation) which carries the risk, irrespective of length of time in theatre. The question is perhaps largely an empirical one but, as was demonstrated in the previous section, the ranking of procedures in terms of risk can change markedly when different exposure measures are used.

Table 8 **Probability of fatality from selected medical procedures**

Drugs per Rx (1974) commonly less than	1×10^{-6}
Vaccinations, 1967–76 against 8 diseases	1×10^{-6}
Anaesthesia in major surgery (1973)	4×10^{-5}
(1951)	6×10^{-4}
Child bearing, risk per maternity (1976)	1×10^{-4}
(1953)	5×10^{-4}
Liver biopsy in 1960s	2×10^{-4}
in 1950s	2×10^{-3}
Treatment with thiouracil 1943–45	4×10^{-3}
X-ray treatment on ankylosing spondylitis	1×10^{-2}
Radium-224 treatment of ankylosing spondylitis	4×10^{-2}
Use of thorotrast as arteriographic contrast medium	6×10^{-2}

Source: Pochin 1981 and 1982

Note: 10^{-6} indicates one in a million or 0.0000001

The second point concerns the time-lag between event and outcome – procedure and fatality. Consideration of time-lag is particularly important when comparing procedures with 'immediate' fatality risk such as anaesthesia in major surgery, with treatments such as x-rays for ankylosing spondylitis where the gap between exposure and death might be quite lengthy. As with occupational risk data, it might be more meaningful to compare risks in terms of life-expectancy lost.

The risk trends over time for a number of procedures are downward. Pochin (1981) for example, details the marked downward trends in fatality risks from anaesthesia, childbirth and abortions during the last twenty years or so. Trends in childbirth and abortion risks are illustrated in Figures 4 and 5. Pochin also notes that the fatality risks of techniques such as liver biopsy have fallen since their initial use as clinical experience is accumulated and practitioners move higher up on their 'learning curves'. In general, the change over time for medical risks follows a predictable pattern of exponential decay. Risk levels tend to decrease at a decreasing rate as some minimum threshold is approached.

A topic of some debate in recent years has been the level of risk associated with prophylactic vaccinations. Clearly, where health care decisions by individuals are elective there will be a concern to compare the probability of contracting the disease without vaccination (with its attendent morbidity and/or mortality), with the probability and severity of any adverse side effects that may result from vaccination. Vaccination is a good example of *risk balancing*, as Jones and Akehurst (1980) note, although in the case of whooping-cough vaccination for example, 'the picture is complicated by the need to compare the risk of convulsions (etc) in the *vaccinated child* with the increased risk *in the population* of epidemics of whooping-cough which could result from having only a low percentage of immune individuals'.

Pochin (1981) examined vaccination fatality risks in the period 1967–76 and observed a fatality rate of 3.3

Figure 4 **Number of maternal deaths per 10^6 maternities (OPCS, 1979)**

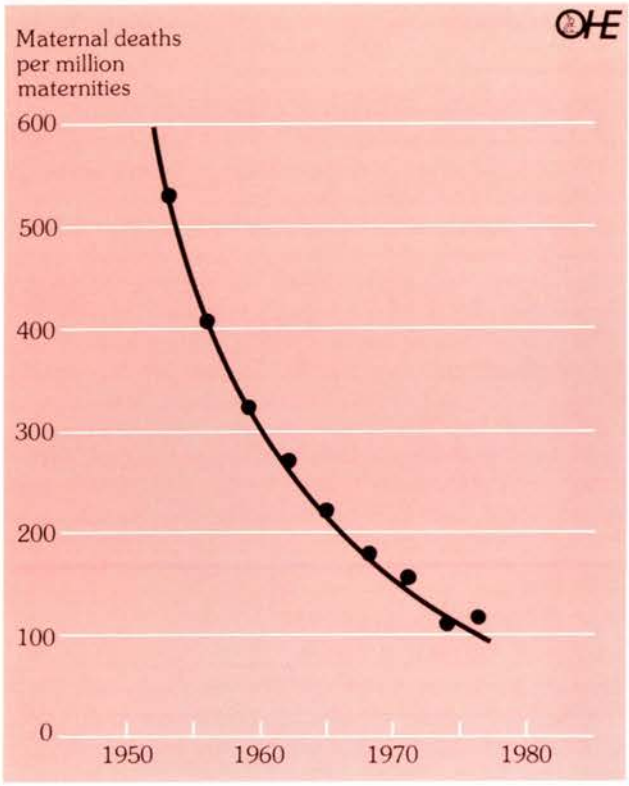


Figure 5 **Number of maternal deaths per 10^4 legal abortions with 90% confidence limits (Lewis, 1980)**

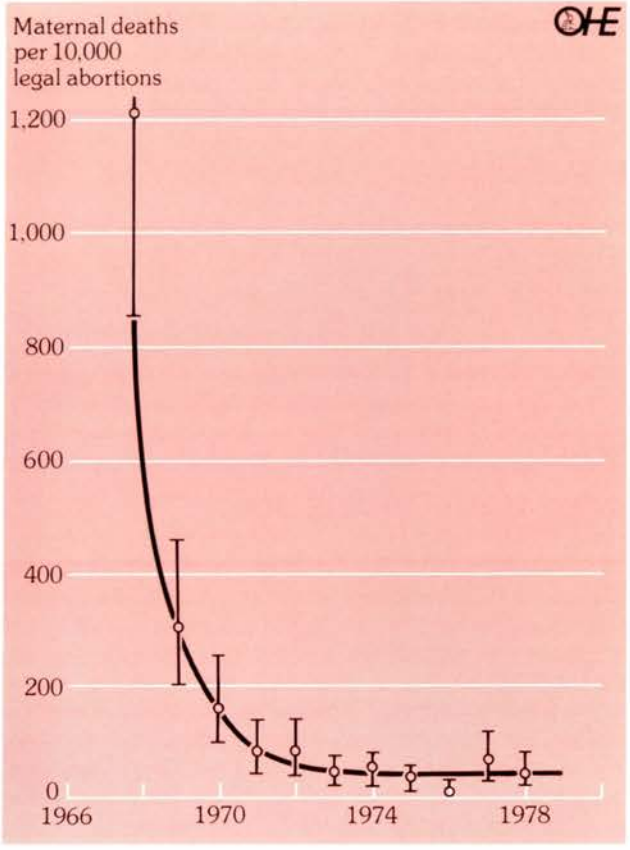


Table 9 **Estimates of risk of brain damage following pertussis (whooping-cough) vaccination**

Researcher	*Adverse outcome*	*Frequency*
Strom (1960)	Vaccine induced encephalopathy	1 per 6,000 injections
	Vaccine induced encephalopathy leading to death	1 per 16,500 injections
Malmgren (1960)	Severe reactions	1 per 50,000 immunisations
Strom (1967)	Destructive encephalopathy	1 per 170,000 immunisations
Dick (1974)	Brain damage	1 per 10,000 immunisations
Prensky (1974)	Severe encephalopathies	1 per 180,000 immunisations
Stewart (1977)	Brain damage and mental defect	Between 1 in 10,000 and 1 in 60,000 immunisations
Grist (1977)	Permanent mental retardation	1 in 135,000 immunisations
Stewart (1979)	Brain damage and mental defect	Between 1 in 17,000 and 1 in 52,000 immunisations
Meade (1981)	Brain damage following any neurological event after vaccination	1 in 155,000 injections
NCES* (1981)	Persistent neurological damage in previously normal children one year after immunisation	1 in 310,000 immunisations or 1 per 100,000 children receiving the full course of three injections

* National Children Encephalopathy Study

Source: Wells (1984)

deaths per year and relating this to an estimate of vaccination activity he gives the fatality risk as being in the order of 1 per million vaccinations similar to that for medicines per prescription.

Yet the reporting of fatality data only gives a partial picture of the risks involved with vaccination. What other morbidity points can be identified on the risk continuum? One of the widely reported side effects of whooping-cough (Pertussis) vaccination in children is that of brain damage. Table 9 details some estimates of adverse outcomes by type and probability. Taking the data estimated by Grist (1977) for example, for every 135,000 immunisations, there is one child who becomes permanently mentally retarded. It is also interesting to note the two estimates of the risk of permanent neurological damage made by the National Childhood Encephalopathy Study (1981); for one injection the probability is 1 in 310,000 but if the child completed the course of three injections the risk compounds itself to 1 in 100,000.

FROM THE AVERAGE TO THE INDIVIDUAL

All the estimates in Pochin's compendium in Table 8 indicate *average* risks, but not all patients exposed to a particular procedure will be 'average' in terms of the statistical distribution of the sample from which the estimate was made.

Thus the likelihood of event 'A' happening may be conditional upon (or related to) characteristic 'B' being present. Depending upon the availability of data it is usually possible to 'sub-set' the risks of a procedure according to patient and other characteristics. In the example of needle biopsy for the liver the risk in the 1960's was estimated as being 2×10^{-4}. But the hazard is due mainly to bleeding from the liver after biopsy and therefore the risk will be higher in distinct groups of patients whose liver disease is associated with defects in blood coagulation or with raised portal vein pressure.

Attempts to quantify relative risk factors associated with treatments vary in sophistication and are often limited due to availability of data. The recent economic evaluation of heart transplantation (Buxton *et al* 1985) contains an example of applying multi-variate statistical analysis to patient survival data in order to determine which patient, donor and treatment characteristics are significant risk factors or predictors of survival. The main results of this analysis were that treatment risks were falling as time progressed (ie survival was improving) and also that the age of the donor was a significant risk factor with poorer survival in cases where donors were relatively older.

This conditional likelihood approach is a way of using epidemiological data and clinical experience to modify the average risk estimate to the individual case. This relates back to the earlier section on subjective probability estimation by the individual clinician attempting to produce a prognosis for a specific patient. The *method* of moving from the general (average) to the specific risk assessment will largely depend upon the availability of data on relative risk factors for patient sub-groups. Where no such data exist or are of dubious validity the clinician's assessment of risk will be more a function of his judgement and experience.

In the following sections the treatment risks associated with pharmaceuticals and surgery will be more closely examined. The aim will be to examine the data which is routinely available with which to assess risk and to investigate its application and reliability.

6a PHARMACEUTICAL RISKS

One of the main features of prescription medicines is that of direct government involvement on grounds of safety regulation. Pharmaceutical risks are assessed prior to marketing and monitored after marketing. Many medicines never reach the market because they do not achieve acceptable safety standards and similarly some medicines are withdrawn from the market following reports of serious or fatal adverse reactions.

The aim of this section is to attempt some assessment of pharmaceutical risks in the context of the recent history of pharmaceutical safety regulation. The focus of attention will be on post-marketing surveillance and the reporting of adverse reactions (ADRs) to the Committee on Safety of Medicines (CSM) under the 'yellow card system'. How well can we measure the incidence of adverse outcomes using these data and what risks are the public exposed to from prescribed medicines?

THE BACKGROUND TO RISK REGULATION

When the effects of thalidomide became known in 1961 there was immediate and obvious public concern regarding the risks of prescribed medicines and the standards of safety regulation. There was an understandable public demand for an increase in consumer protection which led the government to establish the Committee on Safety of Drugs (CSD) in 1963 under the Chairmanship of Sir Derrick Dunlop.

The Dunlop Committee was a panel of experts who reviewed the evidence on new medicines and offered advice on their toxicity to the government. The CSD had no legal power and operated with the voluntary agreement of the pharmaceutical industry. The 1968 Medicines Act radically changed the nature of safety regulation in the UK. In particular, a licensing system was introduced, regulating clinical trials on new pharmaceutical entities and the marketing of new medicines. The government assumed the power and responsibility to decide which medicines were 'safe' to be marketed and which were not.

The evidence on the safety of medicines is currently reviewed by the Committee on Safety of Medicines (CSM) who advise the DHSS whether or not to licence the new medicine. The CSM examines data on the safety, quality and efficacy of medicines both before clinical trial and marketing as well as after marketing.

POST-MARKETING SURVEILLANCE AND ADVERSE DRUG REACTIONS (ADRs)

When a new medicine is licensed for marketing the evidence on its risks will have been assessed from toxicity tests in animals and from the results of clinical trials in humans. The relatively small number of patients (about 1,000–2,000) exposed to a particular medicine before marketing limits the detection of rare adverse reactions (ADRs). Generally the number of patients studied during clinical trials will only be sufficient to detect those ADRs that occur with a relatively high frequency – at least 0.2 per cent to 1 per cent depending upon the spontaneous background incidence of the disease.

The relationship between levels of risk and the minimum number of patients to be studied to detect a given

Table 10 **Minimum number of patients who need to be exposed to a drug to detect a particular level of risk at varying levels of background incidence**

Incidence (risk) of ADR to be detected	*Spontaneous background incidence of the adverse event*	*Minimum number of patients to be exposed*
1 in 100	0	360
	1 in 10,000	520
	1 in 1,000	730
	1 in 100	2,000
1 in 500	0	1,800
	1 in 10,000	3,200
	1 in 1,000	6,700
	1 in 100	35,900
1 in 1,000	0	3,600
	1 in 10,000	7,300
	1 in 1,000	20,300
	1 in 100	136,400
1 in 5,000	0	18,200
	1 in 10,000	67,400
	1 in 1,000	363,000
	1 in 100	3,255,000

Note: For example, to detect an ADR occurring at the rate of 1 in 1,000 (exposed) when the spontaneous background incidence is 1 in 1,000 (in the absence of the drug) will require observations on at least 20,300 patients taking the drug.

These figures were calculated from the normal approximation to the binomial distribution. The spontaneous background incidence was assumed to be known and fixed. A false positive error of 0.05 and a false negative error of 0.05 were assumed.

Source: Committee of Safety of Medicines (1985a)

rate of ADR is presented in Table 10. For example, to detect an ADR which occurs at a rate of 1:1,000 (exposed patients) when the spontaneous (natural) background incidence is also 1:1,000 (in the absence of the medicine) will require observations on at least 20,300 patients taking the medicine.

Given the very large numbers of patients that need to be exposed to a medicine in order to detect less common ADRs, the risk assessment of medicines is unlikely to be 'complete' before the medicine is marketed. There is a need therefore for methods of post-marketing surveillance to allow continued risk assessment as greater patient experience is gained. The CSM currently has three main sources of post-marketing information on ADRs: firstly from reports and articles in the clinical literature; secondly from post-marketing surveillance studies carried out by pharmaceutical companies on a voluntary basis; and thirdly reports of ADRs from medical practitioners in the form of 'yellow cards' sent to CSM.

THE YELLOW CARD SYSTEM

The adverse reaction reporting system was inaugurated on 24 May 1964 by the Dunlop Committee who wrote to all UK doctors and dentists asking for reports of 'any untoward condition in a patient which might be the result of drug treatment'. A Register of Adverse Reactions was established and doctors were provided with a supply of yellow reply-paid postcards for reporting suspected ADRs.

The current guidance given to doctors is that they should report ADRs or suspected ADRs on relatively new products. The British National Formulary (BNF), for example, and other prescribing guides marks the 'newer' products with an inverted triangle ▼ indicating that any suspected ADRs from these medicines should be reported. For these newer medicines, the Joint Formulary Committee (1985) of the BNF recommend that:

> 'Doctors are asked to report any adverse or any unexpected event, however minor, which could conceivably be attributed to the drug. Reports should be made despite uncertainty in the doctors mind about a causal relationship, irrespective of whether the reaction is well recognised, and even if other drugs have been given concurrently.'

The yellow card system is therefore one of voluntary reporting by doctors and dentists to the CSM whenever they have *suspicion* of an ADR. This system of voluntary reporting makes up the vast majority of adverse reactions placed on the Register of Adverse Reactions.

ADRs AND RISK ASSESSMENT

An early study by Girdwood (1974) used data from CSM to examine the fatality risks associated with various medicines in different therapeutic categories. Although the results of this study are now 12 years out of date, Girdwood found that for commonly prescribed medicines the probability of fatality was less than 1 in a million. The measure of risk that Girdwood used was the annual number of fatal ADRs divided by the annual number of prescriptions dispensed. Thus for diazepam (Valium) he calculated that for every million prescription items dispensed there would be an expected 0.76 fatalities.

As a measure of risk there are obvious problems with both the numerator and the denominator. Firstly the numerator, relying heavily on voluntary reporting, is likely to suffer from under-reporting. Griffin and Weber (1985) note that in the period 1972–80, 80 per cent of doctors eligible to do so did not report any ADRs, and that the majority of reports come from the enthusiastic few who have already sent in several reports! The extent to which voluntary reports reflect the true incidence of ADRs is therefore questionable although such data often remain the best estimate. Furthermore, there is the problem of association, causation and the timing of a suspected ADRs; the latency period of some reactions may be such that the link between prescription/consumption and reaction is not obvious.

The second problem is that of the denominator. The number of items *dispensed* is not likely to be the most accurate measure of patient exposure to a particular medicine. If the data were available, two obvious refinements would be to allow for dosage differences between prescriptions and for patient compliance – not all items dispensed will actually be consumed.

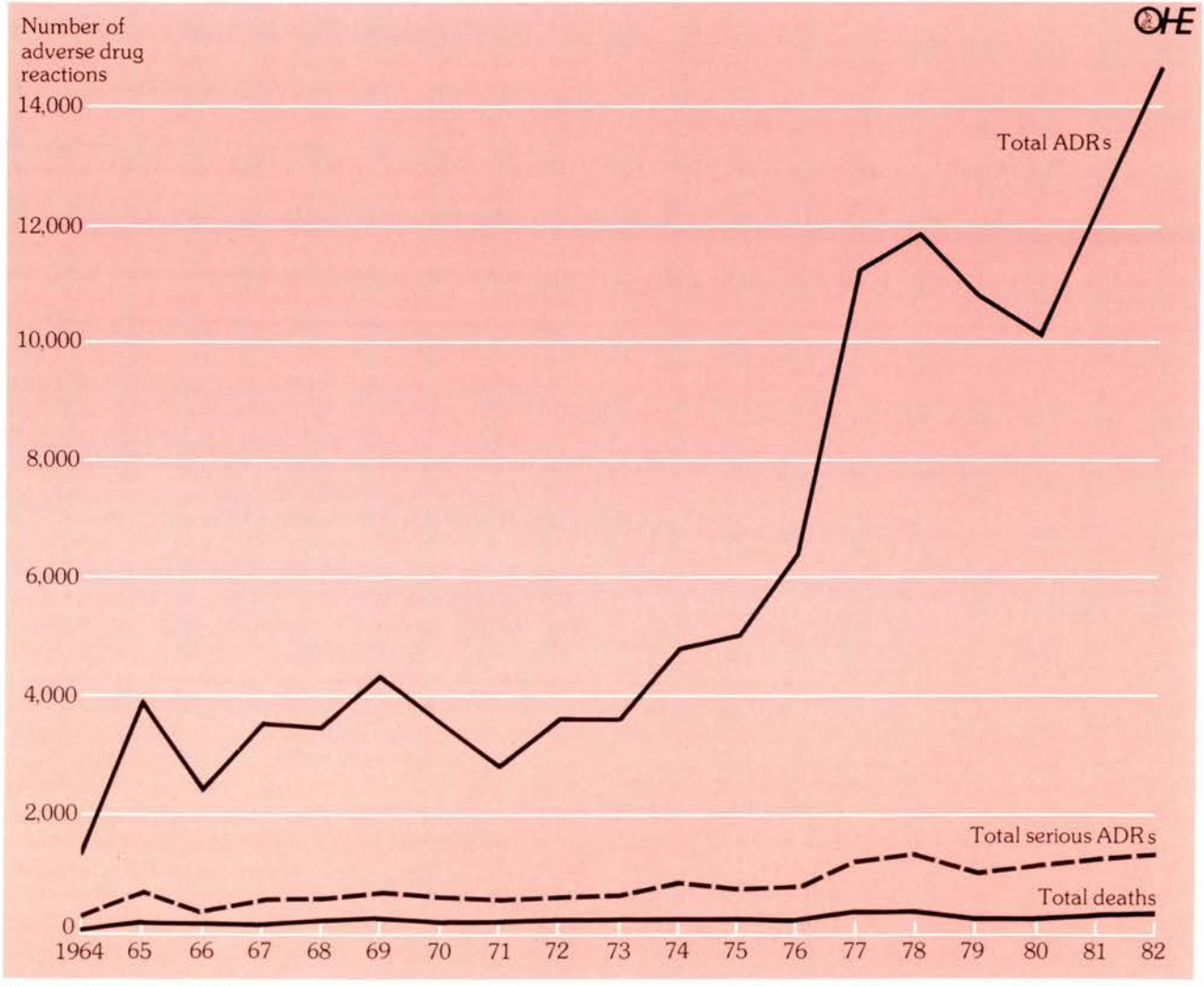
In the knowledge of these caveats about the data, to what extent can ADRs be used as a measure of risk? Table 11 lists the total number of ADRs in the UK in period 1964–1982. Thus in 1982 there were 14,701 total reports of which 9.4 per cent were 'serious' and 2.3 per cent were fatal. As a percentage of total reports in this period, both fatal and serious reports have declined. Trends in reports are presented in Figure 6. Replicating the approach of Girdwood for all prescriptions it can be calculated that the total number of ADRs per million prescriptions dispensed in 1982 was 38.4 and that there were 0.89 fatal ADRs per million items dispensed. During

Table 11 **Spontaneous reporting of total, fatal and serious adverse drug reactions (ADRs) per year, 1964–82, UK. Percentage of deaths and serious ADRs in parenthesis**

Year	*Total reports*	*Total deaths*	%	*Total serious ADRs*	%	*Prescription items dispensed (millions)*	*ADRs per million prescriptions*	*Fatal ADRs per million prescriptions*
1964	1,415	86	(5.9)	279	(19.7)	238.7	5.9	0.36
1965	3,987	169	(4.2)	617	(15.5)	278.9	14.3	0.61
1966	2,386	152	(6.4)	379	(15.9)	299.6	8.0	0.51
1967	3,503	198	(5.7)	511	(14.6)	309.7	11.3	0.64
1968	3,486	213	(6.1)	574	(16.5)	306.3	11.4	0.70
1969	4,306	271	(6.3)	657	(15.3)	302.5	14.2	0.89
1970	3,563	196	(5.5)	600	(16.8)	306.0	11.6	0.64
1971	2,851	203	(7.1)	557	(19.5)	304.5	9.4	0.66
1972	3,638	211	(5.8)	576	(15.8)	315.5	11.5	0.67
1973	3,619	224	(6.2)	617	(17.0)	324.7	11.1	0.69
1974	4,815	275	(5.7)	822	(17.0)	336.9	14.3	0.82
1975	5,052	250	(4.9)	733	(15.3)	346.2	14.6	0.72
1976	6,490	236	(3.6)	882	(13.6)	360.0	17.9	0.65
1977	11,255	352	(3.1)	1,282	(11.4)	363.6	31.0	0.96
1978	11,873	396	(3.3)	1,314	(11.1)	378.1	31.4	1.04
1979	10,881	286	(2.6)	1,096	(10.1)	375.1	29.0	0.76
1980	10,179	297	(2.9)	1,109	(10.9)	374.0	27.2	0.79
1981	12,357	303	(2.5)	1,243	(10.1)	369.9	33.4	0.82
1982	14,701	340	(2.3)	1,382	(9.4)	383.3	38.4	0.89

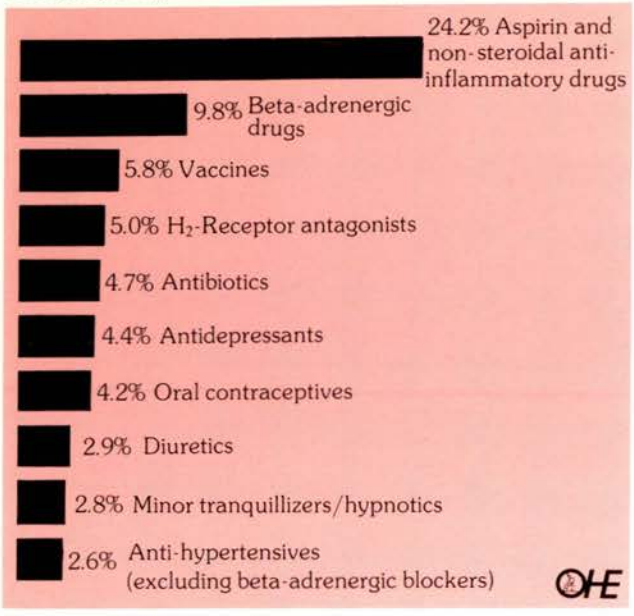
Source: Speirs and Griffin (1984), OHE (1984)

Figure 6 **Spontaneous reporting of total, fatal and serious adverse drug reactions (ADRs) per year 1964–82 (UK)**



Source: Speirs *et al* (1984)

Figure 7 **Adverse reactions reported by therapeutic category July 1976 – June 1982 (total 61,597 reports)**



Source: Data: Speirs *et al* (1984)

this 18 year period the annual risk of fatality from prescribed medicines only once rose above 1 in a million.

Obviously this is an aggregate figure and there are variations around this mean. Particular groups of medicines may be more likely to produce ADRs and certain patient groups (eg the elderly) may be more susceptible to ADRs than other patients.

The 10 therapeutic groups of medicines most commonly giving rise to adverse reactions are presented in Figure 7 for the period 1976–82 and these account for 67.2 per cent of all adverse reactions reported in that period. Non-steroidal anti-inflammatory agents are the most frequently reported class of medicine accounting for 24.2 per cent of all reports in this period.

The logical extension of this approach is to attempt to disaggregate within a given therapeutic class of medicines and compare particular medicines in terms of this risk measure of ADRs per prescription. Speirs (1986) recently completed such an exercise for medicines within the non-steroidal anti-inflammatory (NSAID) group. Many of the medicines in this group are similar in efficacy to analgesics such as aspirin, however they are sometimes preferred to aspirin for the treatment of conditions in elderly patients such as rheumatoid arthritis, osteoarthritis and minor rheumatic conditions. Of particular risk interest in this group is benoxaprofen (Opren) which was withdrawn from the market in 1982 some two years after launch, the CSM having received some 3,500 reports of adverse reactions including 61 deaths.

Figure 8 details the ADR reporting over time associated with medicines from the NSAID group: such as fenbufen, benoxaprofen, tolmetin, oxyphenbutazone and phenylbutazone. The first three of these have been introduced fairly recently while the CSM yellow card system has been in operation, while the latter two are well established and predated the yellow card system. There appears to be two distinct patterns over time. The new medicines have high initial reports of ADRs per million prescriptions which rapidly tail-off while the 'older' medicines have fairly constant, relatively low, levels of reported risk similar to the overall average of 38 ADRs per million prescriptions calculated in Table 11 previously. This marked decline in reporting raises questions about how indicative the early reports are of the actual incidence of ADRs. Doctors clearly report far more frequently on new medicines than on established medicines, but to a large extent this may not reflect differential risk but simply that they have been *specifically asked* to report their suspicions on any new medicines.

A further interesting aspect of Figure 8 is the risk relationship between NSAIDS which have been withdrawn from the market and those which remain. Over the last few years several of the NSAIDs have been withdrawn and they include: benoxaprofen (Opren), feclofenac (Flenac), feprazone (Methrazone), flufenanic acid (Meralen), oxyphenbutazone (Tandacote, Tanderil) and zomepirac (Zomax). In addition the CSM has recommended that the use of phenylbutazone be 'restricted'.

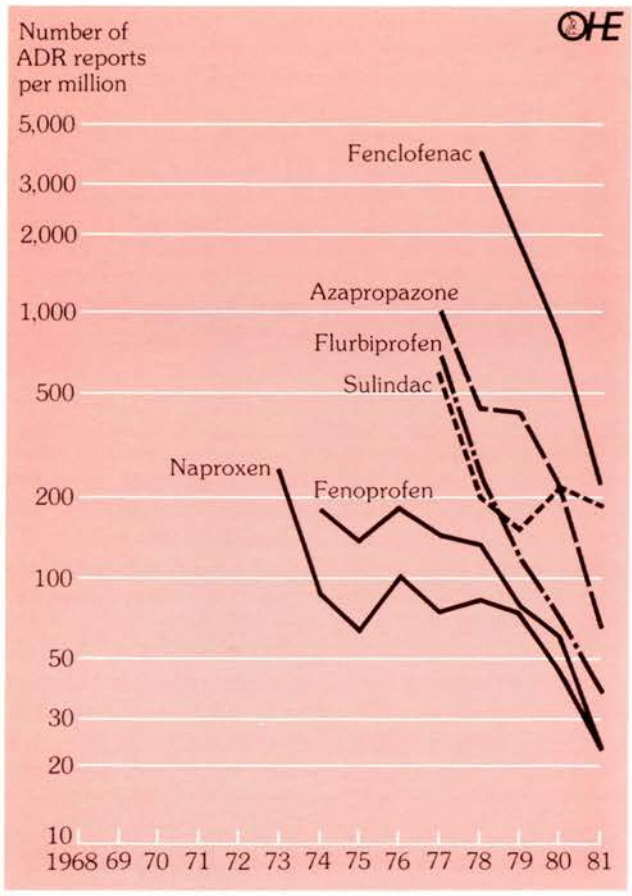
Assuming for the moment that the efficacy or patient

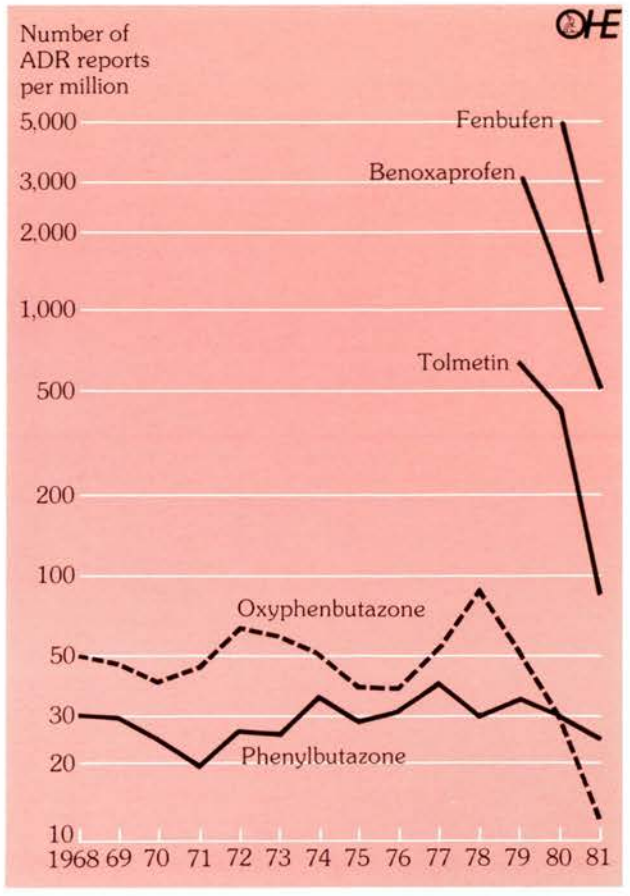
Table 12 **Relative risk of fatal blood dyscrasias as judged by CSM using yellow card reports and estimated number of prescriptions**

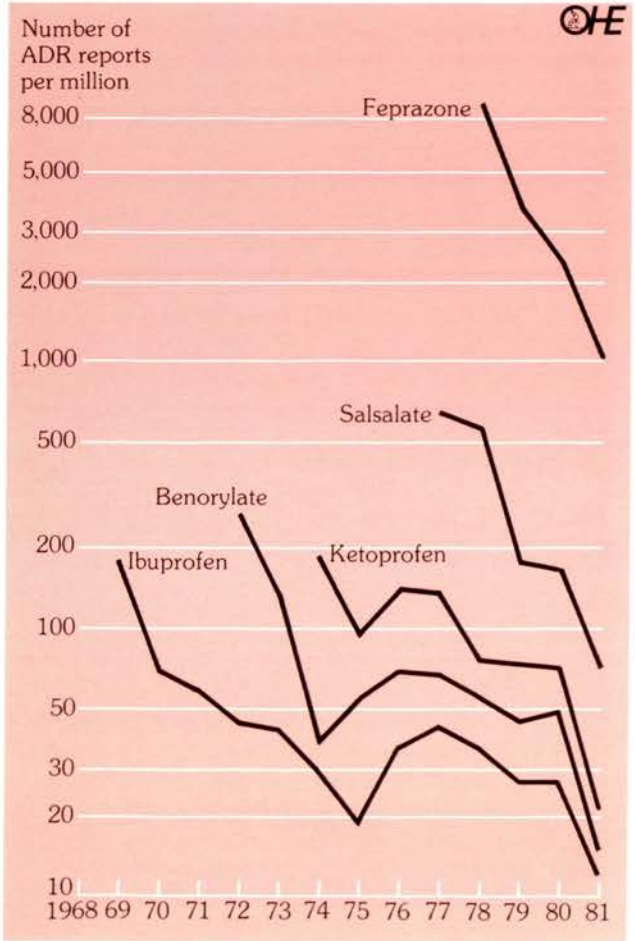
	Total number of reports of blood dyscrasia	*Number of reported deaths due to dyscrasia*	*Index of risk fatality*
Non-steroidal anti-inflammatory agents in common use:			
Ibuprofen	97	12	1
Indsmethacin	135	33	2
Piroxicam	37	5	2
Naproxen	77	15	3
Fenoprofen	31	7	5
Restricted or withdrawn medicines:			
Phenylbutazone	531	301	20
Benoxaprofen	72	13	26
Oxyphebutazone	195	109	43
Disease modifying medicines:			
Penicillamine	94	20	70
Gold	56	23	179

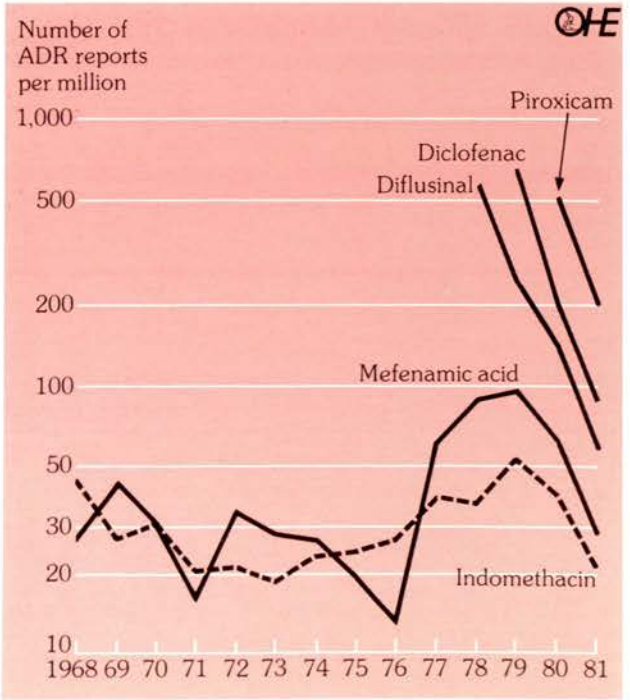
Source: Committee on Safety of Medicines Update. BMJ (1985b)

Figure 8 **Adverse reaction reports per million prescriptions for selected non-steroidal anti-inflammatory drugs (1968–1981)**









Source: Speirs (1986)

benefit of all the agents presented in Figure 8 is fairly similar and that the decision to withdraw a medicine is made solely on the basis of relative risk, it is not immediately clear why oxyphenbutazone (roughly 50 ADRs per million scripts) has been withdrawn and yet other apparantly 'more risky' compounds remain.

The CSMs calculations of fatality risks due to blood dyscrasias from prescribed medicines in the NSAID are presented in Table 12. The 'index of risk of fatality' being calculated as the 'ratio of reported fatalities per number of prescriptions for a medicine to the number of reported fatalities per number of prescriptions for ibuprofen'. In short it calculates risks relative to one NSAID, ibuprofen, using fatal blood dyscrasias ADRs as the numerator and estimated scripts dispensed as the denominator. This table gives some indication for the fatal risk relativities between these medicines which have been withdrawn and those which remain. But CSM stress caution when interpreting these data, stating that:

> 'While the difference between naproxen and phenylbutazone, for example, is uncontroversial, we cannot say with any confidence that fenoprofen is two and a half times more likely to cause a fatal blood dyscrasia than piroxicam. The number of reports for both medicines are too small for any firm conclusions about the relative risk to be drawn . . .'

Depending upon how the various statistics are presented there seems a danger that simple quantitative methods will not cope adequately with the many qualitative aspects concerning ADRs as a risk indicator. At one level there is the obvious concern that small numbers of observations cannot yield reliable statistics with which to calculate relative risks. A more subtle concern though is the relationship between fatal and non-fatal ADRs as perceived by those responsible for assessing the 'riskiness' of medicines.

Clearly, *all* ADRs will adversely influence patients' health, some reducing quality of life and some proving fatal. It is not immediately clear from CSM reports what weight is to be attached to fatal outcomes versus non-fatal. To express this another way, how many serious but non-fatal ADRs are equivalent to one fatal ADR? The problem is essentially one of determining relative values with which to trade-off quantity and quality of life aspects.

Such judgements are important when comparing two or more medicines of similar efficacy in terms of risk. If new product A is associated with 5,000 serious but non-fatal ADRs per million prescriptions and new product B has only 50 ADRs per million, but 5 of these are fatal, which of the two medicines can be judged to be the more risky? Applying the principles of expectation discussed earlier it might be possible to calculate the expected *outcomes* for the two medicines. But the choice of strategy can only be evaluated when *values* have been assigned to the health outcomes. *Implicitly* some values are already being applied by risk regulators because decisions *are* made concerning risk on the basis of ADR data.

In terms of decisions about which medicines to withdraw on grounds of risk the crucial factor is that judgements be made on the basis of expected costs (risks) and expected benefits. Both sides of this coin will have probabilistic health outcomes which must be traded-off against probable death or disability. The problem is one of determining the expected value or utility of one decision strategy – withdrawing the medicine – against its alternative. The question of relative values in this context is further discussed using the example of Opren at the end of *Section 9*.

PRESCRIPTION EVENT MONITORING (PEM) AND RISK ASSESSMENT

The main problem with monitoring methods such as the yellow card system is that of under-reporting. The extent of this under-reporting of ADRs is unknown and probably varies depending upon the nature of the medicine and reaction. If the level of under-reporting were similar between medicines then the yellow card system might be a more useful indicator of the relative risk of different medicines (eg ADRs per million scripts).

A prime concern, however, is that reporting rates are influenced by publicity surrounding a particular medicine. The data indicate that whenever there is publicity about problems with a particular medicine then the reporting rate for that medicine usually rises. Again, this highlights the important role of publicity and the media influencing perceptions of relative risk held by the public and prescribing doctors.

The problem of yellow card under-reporting was highlighted when the problems associated with the medicine practolol became known in 1974. Practolol (~~Inderal~~ Eraldin) is a beta-blocker which was prescribed for such indications as relief of angina and high blood pressure. It was marketed in 1970 after stringent pre-marketing tests and in the next four years there had accumulated some 200,000 patient-years of experience with the medicine. The marketing company, ICI, commented that '. . . (then) came the bolt out of the blue and we learnt that it could produce in a small proportion of patients a most bizarre syndrome, which could embrace the skin, eyes, inner ear, and the peritoneal cavity' (quoted in Laurence and Black, 1978).

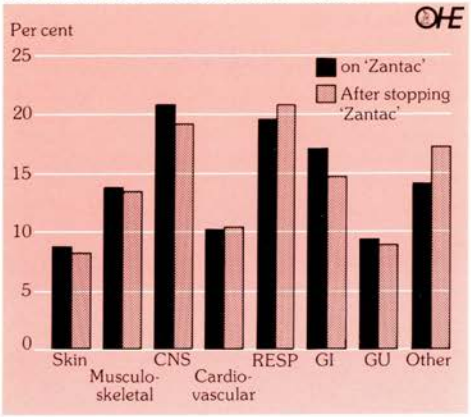
The concern following practolol was that the yellow card system had failed to detect the rare adverse reaction early enough and a number of new techniques were suggested for monitoring. A change of approach for post-marketing surveillance was the idea of using data on prescriptions dispensed as the starting point for monitoring patients exposed to particular medicines.

Prescription Event Monitoring (PEM) was developed by Professor Inman at the University of Southampton's Drug Surveillance Research Unit. The basic idea of PEM is that copies of prescriptions dispensed for particular medicines are obtained from the Prescription Pricing Authority (PPA) thus enabling researchers to contact prescribing doctors and attempt to identify a cohort of patients who are, or have been, exposed to a given medicine. Prescribing doctors are contacted and asked to provide data on medical 'events' which have occurred to patients exposed to particular medicines.

One advantage of PEM is that the emphasis is on the collection of data on all medical events for given patients, rather than doctors' suspicions that events are related to the medicine. The question of a casual link between ADR and medicine can be investigated by the epidemiologist and does not rely on the opinion of the GP. The second major advantage of PEM is that the system provides a denominator which allows a much better estimate of risk to be made because far more precise data on exposure to the medicine by patient are analysed.

The presentation of data on medical events is made in the form of an 'event profile', an example of which is presented in Figure 9 for the medicine Zantac (ranitidine).

Figure 9 **Profile of various classes of event, expressed as a percentage of all events recorded during and after treatment with 'Zantac'**



Source: Drug Surveillance Research Unit (1983)

this being based on a pilot study of the medicine by the surveillance unit in Southampton. (Zantac follows on from Tagamet (cimetidine) as a medicine prescribed mainly for the treatment of peptic ulcers). The histogram in Figure 9 represents patterns of events which occurred to the patient cohort both during and after treatment with Zantac. Events are presented in terms of the particular body system of origin (eg musculo-skeletal, central nervous system, etc.) The illustration is based on 3,822 events recorded during treatment with Zantac compared with 3,645 events after the patient had stopped treatment, and the comparison indicates very little difference between the two groups of patients. This sort of event profile can also be disaggregated to examine events for particular patient groups (eg the elderly) within a given cohort and therefore offers a good deal of flexibility to the analyst examining such profiles looking for ADR links.

The main drawback with PEM is that, to date, the PPA have only been able to identify and monitor prescriptions for four medicines at any one point in time (although this is shortly set to improve with the computerisation of the PPA). Furthermore the handling and interpreting of report forms is a relatively costly exercise which rises with the size of the cohorts being monitored – cohorts are currently limited to around 10,000 patients. In addition, the experience of PEM has been that only 60 per cent of doctors comply with requests to complete forms, and again it seems likely that percentage compliance will fall as greater demands are made on doctor time with larger patient cohorts across an increasing number of medicines.

In summary it would appear that the treatment risks associated with prescription medicines are small (1 in a million on average) compared with many other risks that individuals accept. A glance at travel risks will confirm that the average patient is more likely to be killed in a traffic accident if she drives to the doctors surgery than by the medicine that the doctor might prescribe. Yet this is no argument for complacency in risk assessment. The problem however, is that as the frequency of ADRs becomes more rare so too does the difficulty and cost of detecting these adverse effects.

6b SURGICAL RISKS

Most people undergo some kind of surgery in their lifetime; be it teeth, tonsils or transplants. The spectrum of surgical complexity is vast and expanding. New technologies and techniques are constantly being developed which render operable the ailments which were previously inoperable. But there are risks. The common currency of surgical risk is either that of the probability of operative mortality (dying in theatre) or of mortality within some specified period post-operatively, often until discharge from hospital.

The aim of this section is twofold. Firstly to examine the issue of diagnostic and prognostic accuracy in a world of uncertainty. The example is taken from de Dombal's (1974) work on suspected appendicitis and decisions on whether to operate or not. Secondly, to present data from the Hospital Inpatient Enquiry (HIPE) on case fatality rates for selected surgical procedures as a source of some crude but routinely available surgical risk estimates.

'TO OPERATE OR NOT TO OPERATE...'

A discussion of diagnostic accuracy could preface a review of any area of medicine or surgery. Diagnosis and prognosis are informed gambles. Hopefully the doctor will be right more times than he is wrong. But in surgery a false diagnosis might result in a patient being exposed to the risk of surgery unnecessarily. This section explores the way in which the diagnostic lottery can be analysed.

A patient presents with abdominal pain. On the basis of signs and symptoms the doctor may or may not diagnose appendicitis. The patient may (in fact) have appendicitis. If the diagnosis is positive they will operate and soon find out if the diagnosis was accurate. If the diagnosis is negative they will not operate. Again this diagnosis may be accurate or not – time will tell.

There are four possible outcomes to this problem: **True-positive** (patient diagnosed as having appendicitis and does have it); **True-negative** (patient diagnosed as not having it and does not in fact); **False-positive** (patient told he does have appendicitis but in fact does not) and **False-negative** (patient told he does not have it but in fact he does). The possible outcomes and errors are presented in matrix form in Figure 10.

In de Dombal's (1974) study a trial was conducted to compare the performance of human and computer-aided diagnosis for abdominal pain. In many respects the computer aid might be said to be an early form of 'expert system'. The doctor feeds diagnostic information into the machine for a particular patient and, given the data it has available to it, the system predicts the likelihood of appendicitis. During the trial the doctors were, in a sense, 'competing' with the machine to see who performed best in terms of diagnostic accuracy.

The results of the trial have been summarised in Figure 11 which is taken from Collinson and Dowie's (1980) discussion of the results. Before the trial the position was D1 – 60 per cent were true-positives (therefore 40 per cent were false negatives) and 25 per cent were false positives. In other words, a quarter of the patients were undergoing surgery unnecessarily because they were diagnosed as having appendicitis but in fact did not have it. During the trial (D2) the true-positive rate rose to 93 per cent and

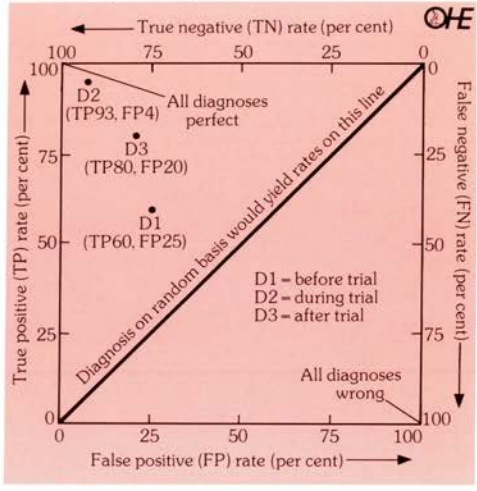
Figure 10 **Decision matrix for diagnosis of abdominal pain**

Diagnosis of doctor	Actual condition of patient	
	Patient has appendicitis	Patient does not have appendicitis
Diagnose appendicitis (operate)	True-positive (TP) rate (93%)	False-positive (FP) rate (4%) (unnecessary surgery laparotomy)
Doesn't diagnose appendicitis (wait)	False-negative (FN) rate (7%) (Delayed detection; appendix may perforate or form abcess)	True-negative (TN) rate (96%)
	100%	100%

Note: TP and FP rates relate to diagnostic performance *during* the trial, point D2 on Figure (11).

Source: based on Collinson and Dowie (1986) after de Dombal *et al* (1974).

Figure 11 **Probability and diagnosis: de Dombal's abdominal pain study**



Source: Collinson and Dowie (1980) after de Dombal *et al* (1974).

false-positives fell to only 4 per cent. After the trial had finished however, diagnostic performance fell back to position D3 on Figure 11.

The dotted diagonal line in Figure 11 illustrates very clearly the nature of the diagnosis gamble, because this line represents the result that would be achieved by some random assignment such as tossing a coin. In other words, a patient with appendicitis has a 50/50 chance of being diagnosed positively or negatively! The proximity of the pre-trial diagnostic performance (D1) to this line is a little worrying. It is interesting to speculate whether, in certain conditions, we would accept diagnosis on a random basis if it performed, on average, better than the doctor opinion (ie higher true (+) and lower false (+) rates).

de Dombal and colleagues could not determine why diagnostic accuracy among the clinicians improved during the trial and then fell again afterwards. One thought put forward (but not endorsed) by them does have some appeal:

> 'One could postulate, for example, that the computer represented an "opponent" to be beaten in a diagnostic "contest" – but this implies that the clinician is so apathetic normally that he requires some such spur to function effectively, and this is an assertion which we do not make.' (de Dombal, 1974, p 379.)

The task of *evaluating* the performance of different diagnostic or screening tests cannot be done without placing values on outcomes. In this example, of particular interest is the value to be placed on the two types of error. A false (+) error results in exploratory abdominal surgery (laparotomy) so the patient is exposed to the risks of anasthesia and surgery along with the other costs such as scarring. The false (–) outcome results in delayed detection of appendicitis which in turn means that the appendix may perforate or form an abcess. To decide which of these is worst is to invoke a value judgement. The evaluation of diagnostic test performance or other screening tests is therefore a combination of estimating outcome probabilities but also of determining the *relative values* associated with these outcomes. (For further discussion on such issues in the context of screening programmes see Simpson, Chamberlain and Gravelle (1978).)

ESTIMATING SURGICAL RISKS

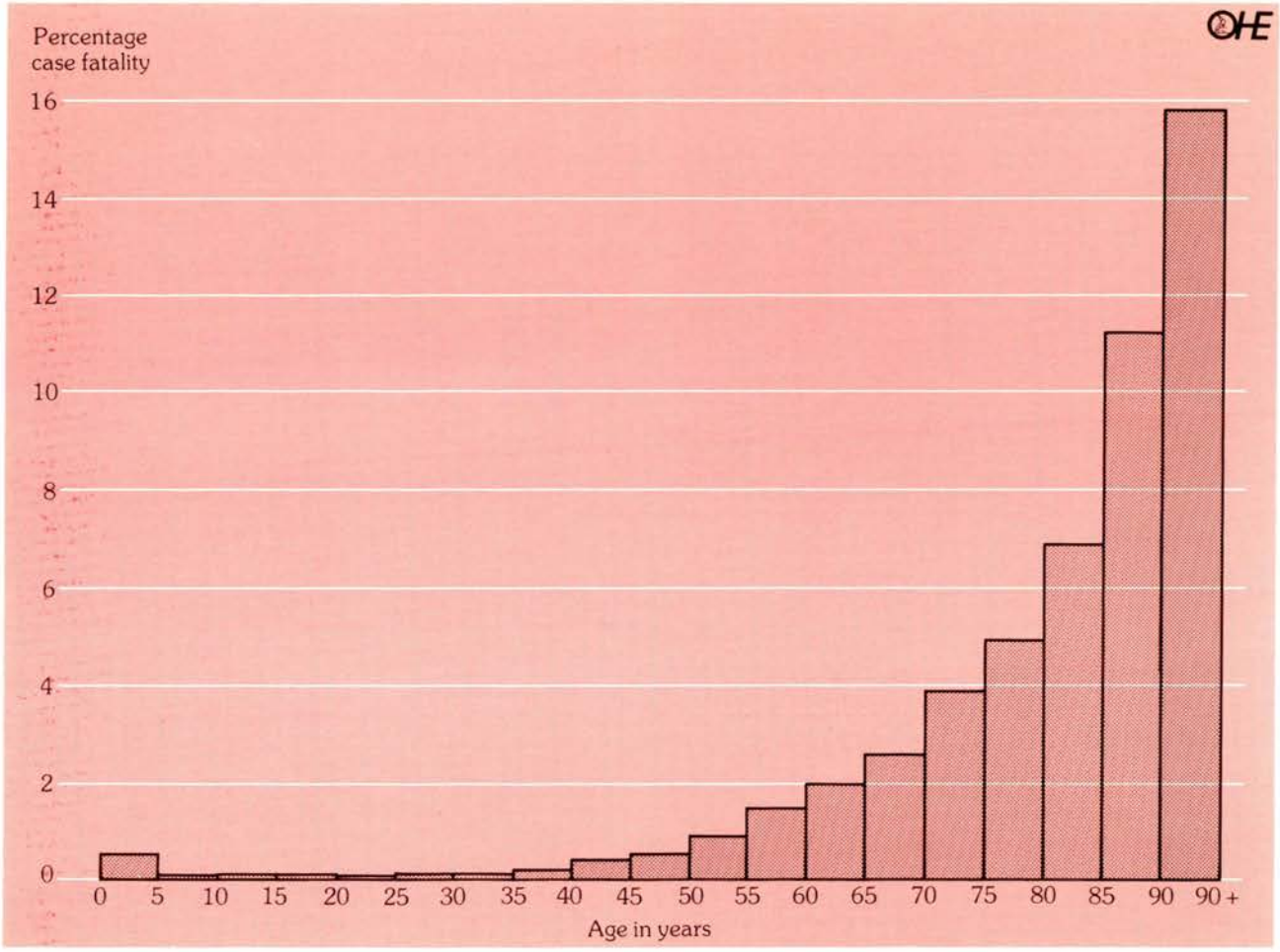
Routine data are available on a sample basis from the Hospital Inpatient Enquiry (HIPE), which details percentage case fatality rates by type of operation or various patient groups. HIPE is based on a one in ten sample of patients and therefore its statistics should be seen as indicative rather than definitive. The definitions and codings for operations are those used by the Office of Population Censuses and Surveys (OPCS), thus 'surgical operations and procedures' are:

> 'any therapeutic or major diagnostic procedure which involves the use of instruments or the manipulation of parts of the body and generally takes place under operating theatre conditions. Where more than one operation is performed during a period of stay, priority is given to the procedure most closely related to the principal condition for which the patient was treated or investigated.'
>
> (OPCS, 1985)

The measure of risk reported here is that of *percentage case fatality* for any given operation. In this context the definition of the 'case' is the period between operation and discharge which therefore provides only a fairly short-term inpatient basis for comparing treatment risks. In addition there are coding problems of attributing cause of death to the operation (eg the patient who dies from myocardial infarction while recovering from appendectomy). Nonetheless the HIPE data indicate the broad dimensions of operative risk.

In 1983 there were approximately 2.2 million surgical operations and procedures performed in England; a rate of about 5 per hundred population. Not surprisingly the fatality risks of surgical intervention increase with age. Figure 12 clearly demonstrates the exponential growth in

Figure 12 **All surgical operations and procedures, percent, case fatality by age. (Males and females, England 1983)**



Source: HIPE.

the case fatality rate with age. This simple aggregate illustration suggests that the probability of being a post-operative inpatient fatality increases sixfold in the thirty years after the age of sixty. The fact that octagenarians withstand surgery less well than adolescents is reflected in the distribution of surgical activity by age which indicates a rapid decline in the number of operations performed in patients aged 75 or greater Figure 13.

Recent trends over time for all operative procedures performed suggest an overall decline – the aggregate picture from HIPE in 1983 is a 1.6 per cent case fatality rate over all procedures. Obviously, because surgical risks vary between operations there is a need to disaggregate the picture over time.

In Figure 14 the recent experience over time is illustrated for five types of surgical intervention. The selection of each type or area of surgical intervention has been chosen to illustrate the range of case fatality rates and hence indicate various *levels* of surgical risk.

Operations on the Skull and Brain (OPCS code 001) had a case fatality rate in 1983 of 7.4 per cent; there appears to be no obvious downward trend in his rate compared for example with **Operations on the Heart and Intrathoracic Vessels** (032) which demonstrate a marked and continuous downward trend in case fatality from 4.9 per cent in 1979 to 3.0 per cent in 1983. Indeed there exists additional evidence for the specific operation of coronary artery bypass grafting (CABG) from English *et al* (1984) that case fatality rates have halved in the period 1977–82, see Table 13.

Table 13 **Risk of early* death following coronary artery bypass graft, UK 1977–82**

	Number of procedures	*Percent mortality*	*Fatality risk per procedure*
1977	2,297	6.4	0.064
1978	2,653	5.1	0.051
1979	2,918	6.1	0.061
1980	4,057	3.7	0.037
1981	5,130	3.7	0.037
1982	6,008	3.2	0.032

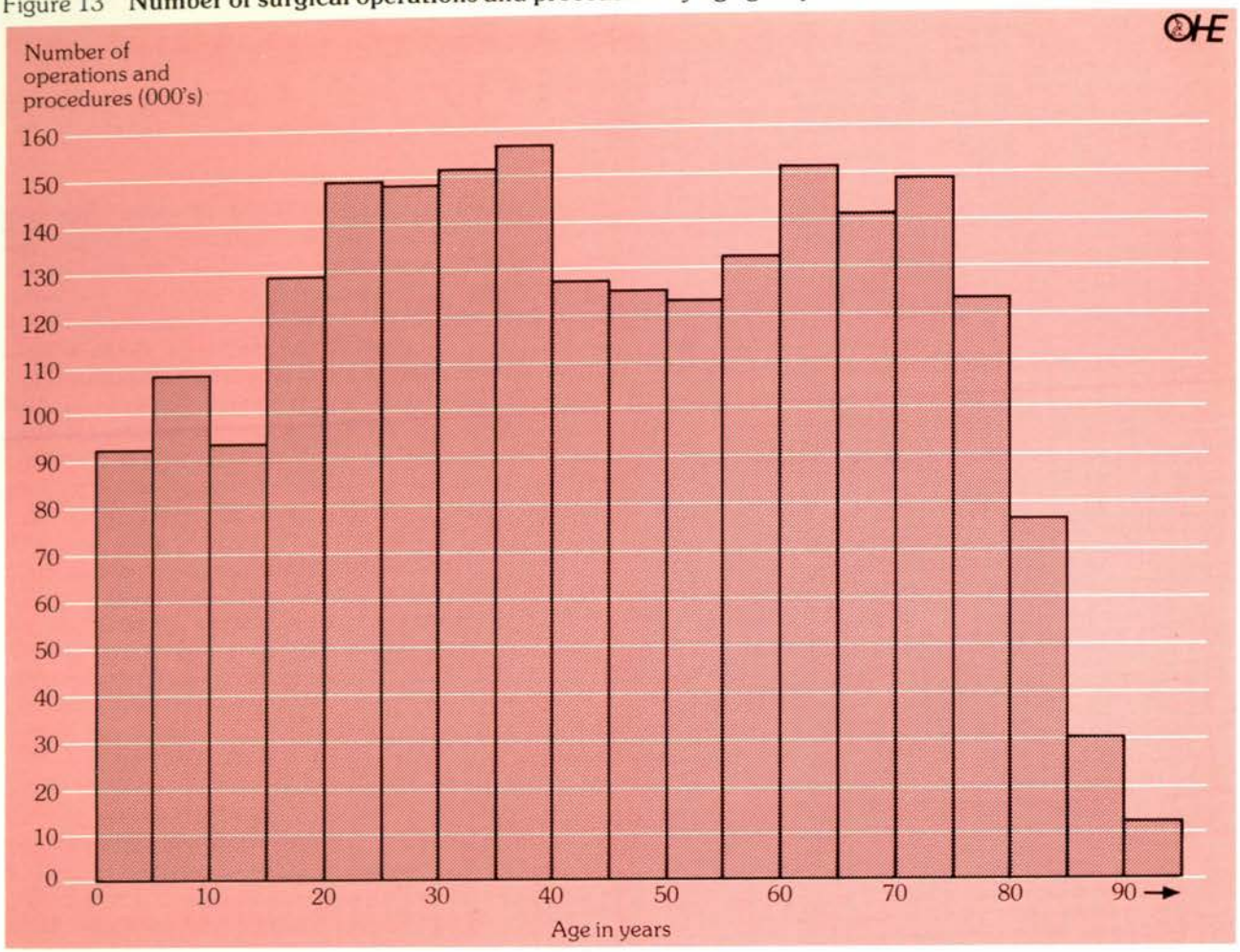
*Early death defined as being within 30 days from operation.

Source: English, Bailey, Dark and Williams (1984)

Some of the baseline markers in Figure 14 are from the many routine and often elective surgical procedures. **Operations on Tonsils and Adenoids** (023) carry a case fatality rate of about 0.01 per cent or 1 in 10,000. Similarly, **Operations on Haemorrhoids** (051) carry a case fatality risk of about 1 in 1,000 (although in 1983 there were actually no fatalities with the 10 per cent HIPE sample of 1,242 patients).

Clearly, as with many other endeavours, there will be diminishing returns in reducing risks of surgery. Heart surgery, for example, appears to be becoming rapidly less risky (in terms of case fatality) as experience and technology advances techniques and patient care. But the rate-of-change of decline is itself falling. As illustrated by

Figure 13 **Number of surgical operations and procedures by age group. (Males and females, England 1983).**



Source: HIPE.

Figures 4 and 5 the observed reductions in risk over time display a typical exponential decay when expressed graphically. A number of procedures appear to have reached a minimum risk level which is remaining relatively constant over time. In Figure 14 both **Operations on the Appendix** (044) and **Tonsils and Adenoids** (023) might be placed in this latter category.

Obviously one of the main factors influencing the reduction of operative risks is the risk associated with anaesthesia itself. Complications and fatality rates attributable to anaesthesia have fallen quite dramatically in the last 30 years – this fall is illustrated in Figure 15. Thus in 1961 there were approximately 21 anaesthetic related deaths per 100,000 operations; falling to 7 per 100,000 in 1967 and slightly less than 4 per 100,000 in 1973.

Although the data are far from ideal it is possible to go some way toward the quantification of treatment risks in both medicines and surgery. Routine surgical risk data are generally in the form of survival data or case fatality rates, but this reflects the fact that the indications for surgery generally carry a greater risk than for other treatment interventions. Pharmaceutical risk data is routinely available on a wider range of health impairments although this too is focussed heavily towards fatality probabilities.

Clinicians can use these 'average' risk data and tailor such estimates into conditional likelihood data for specific patients or 'groups' of patients. But to what extent is the publics' *perception* of risks the same as the professionals'

Figure 14 **Percentage case fatality for selected operation groups (OPCS code) Males and females; England and Wales.**

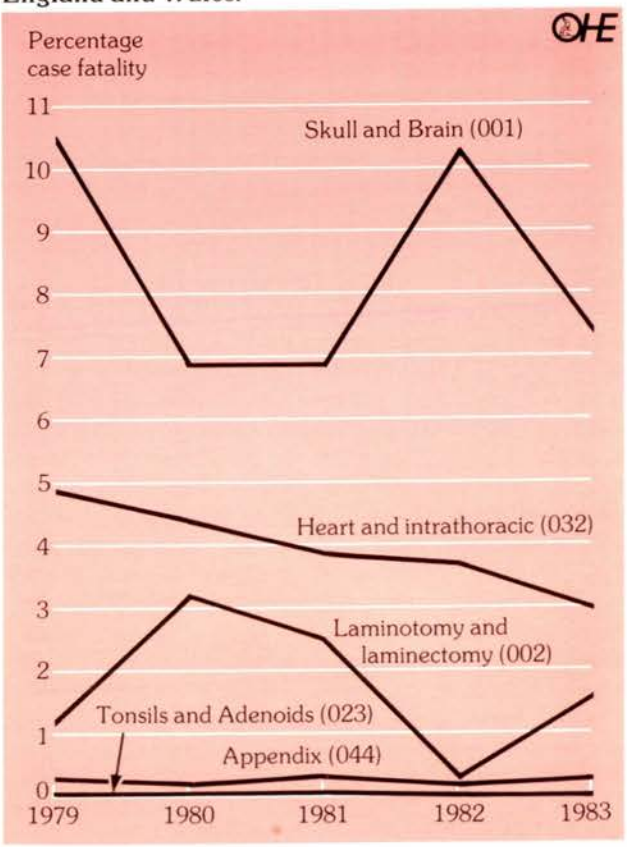
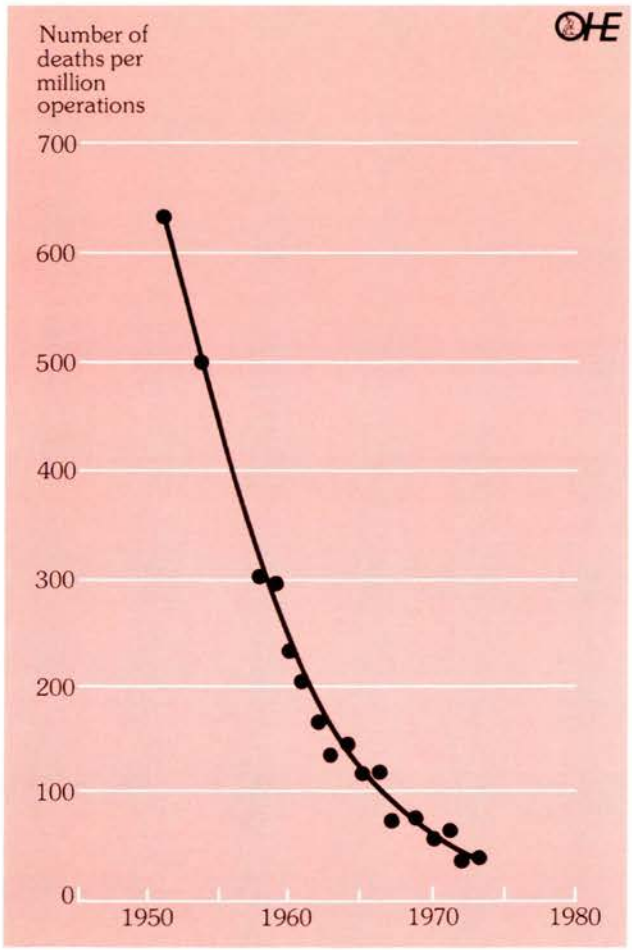


Figure 15 **Number of deaths attributed to anaesthetic per 10^6 operations**



Source: Office of Health Economics (1976).

estimate? If the aim is to incorporate patients' values regarding risks and risk taking into clinical decision making, it is first necessary to investigate the way in which the public may mis-perceive risks in general, and treatment risks in particular.

7 RISK PERCEPTION: A TALE OF BIAS AND ILLUSION

Despite attempts to refine methods of producing 'objective' estimates of risk, individual decision-making will be based on how risks and benefits are *perceived*. The level of risk assessed by the 'experts' in any given situation may only be vaguely related to the level of risk perceived by individuals.

The perception of risk is like the perception of noise, heat or health. Mankind has developed objective measures with which to quantify these phenomena such as decibels, temperature scales or bodily functioning. But it is a person's own perception of the world around him which produces judgements concerning which sounds are louder than others, which climates are warmer than others or which aspects of physical or mental functioning constitute greater or lesser health status. Similarly with risk, the data on probability and outcome can be put before different audiences who will disagree about the 'riskiness' of options because perceptions regarding likelihood will vary along with the values attached to different types of outcome.

The relationship between the 'expert's' risk estimate and the individual's risk perception will be variable and possibly vague. But psychologists have studied the cognitive processes which underly risk perception in an attempt to explore whether there are any *systematic* ways in which individuals 'misperceive' risks. The aim of this section is to review these elements of bias in perception and relate them to the area of clinical risks.

INFORMATION AND JUDGED LETHALITY: 'I READ IT IN THE SUNDAY PAPERS'

Tversky and Kahneman (1974) have studied in detail the way in which individuals introduce heuristics and biases when faced with judgements about risk. One particular influence has been labelled the **availability** heuristic. Stated simply this means that an individual will judge an event to be more likely if instances are easy to recall or imagine. In other words it is the *availability* of 'information' on a given risk topic which influences perception. This information can be in the form of personal experience or more generally in the form of media coverage.

Personal experience is a powerful influence on perception. The person who has lost family or a friend with leukemia is likely to over-state the incidence of this disease in the community. Press and television reports of airline disasters will increase public perceptions regarding the likelihood of death while travelling by plane. The greater the degree of public discussion of low-probability hazards (eg nuclear core melt down) the more *imaginable* such events become in the public mind.

One direct route for examining public risk perception regarding health care and other activities is to ask them questions like: What do you think is the probability of a fatal adverse reaction from a prescription medicine? How many people die annually from asbestos-related diseases? How does the wearing of a seat belt affect your probability of living through the year? The responses can be compared with the 'expert' assessments and discrepancies interpreted as the degree of bias or misperception. Such an approach was adopted by Lichtenstein *et al* (1978) who asked a sample of the public to judge the frequency of 41 causes of death. As a point of reference, respondents were told the annual death toll in the USA for motor vehicle accidents was 50,000 and then they were asked to judge the frequency of the other 40 causes.

Figure 16 presents the results from this study and uses a logarithmic scale to compare the judged number of deaths with the 'official' estimates from public health statistics. If the frequency judgements had been 'accurate' all points would have been on the simple diagonal through the origin. The points on Figure 16 are dotted around this diagonal and Lichtenstein and colleagues plot a curved line through the points to indicate the general relationship that they observed, namely that rare causes of death were over-estimated and common causes of death were under-estimated.

While deaths from stroke and stomach cancer were under-estimated, deaths from pregnancy and smallpox vaccination (which are relatively rare) were over-estimated. Accidents were judged to cause as many deaths as diseases, whereas diseases actually take about 15 times as many lives. In general, and in keeping with the availability heuristic, the over-estimated items tend to be sensational and dramatic (accidents, floods, tornados), whereas the under-estimated causes are more common, less dramatic events (stroke, diabetes).

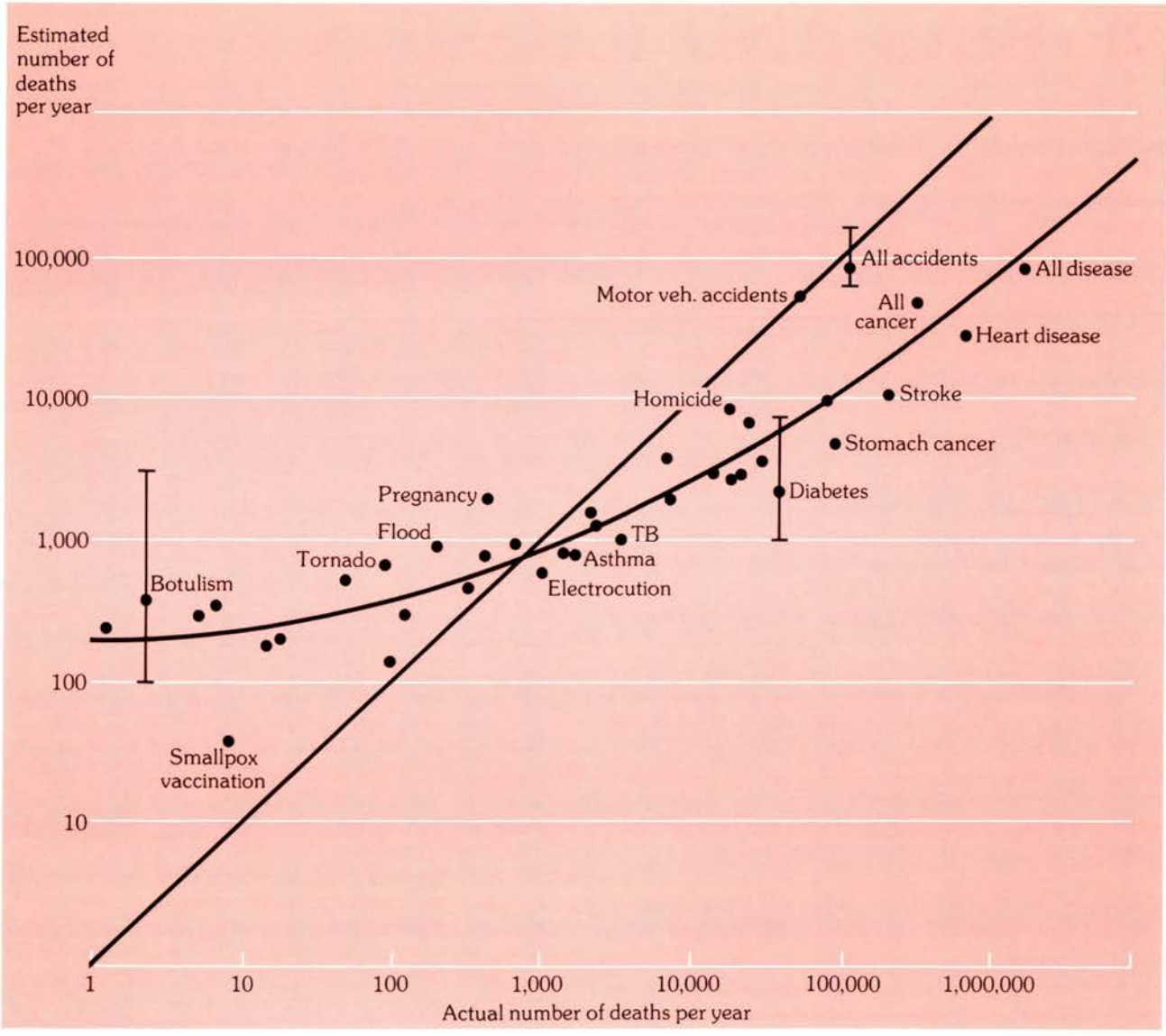
Public over-estimation regarding the likelihood of rare events and disasters is obviously related in part to the information and news generally available to the public. Combs and Slovic (1979) examined this question and reported the (not surprising) fact that newspaper coverage tends to be biased towards the more sensational fatal and life threatening events. Accidents, violence and disasters sell more newspapers than heart disease and stomach cancer. In 1986, it would be interesting to examine the public's perception of the risk of AIDS, which has had extensive and sometimes sensational press coverage, relative to other causes of death.

'IT'S NOT WHAT YOU SAY, IT'S THE WAY THAT YOU SAY IT'

In the earlier discussion of probability it was noted that there is a good deal of ambiguity and inconsistency regarding the public's translation of verbal statements of likelihood into numeric values for probabilities. In addition to this there may be systematic bias in decision making under uncertainty due to the way in which the expected outcomes have been expressed. More precisely, it may depend on whether outcomes are expressed in terms of losses or gains.

An interesting example of this phenomena in a clinical decision making context can be found in McNeil *et al* (1982). The aim of this study was to examine preferences for alternative therapies for the treatment of lung cancer; the main choice being radiation therapy versus surgery. Groups of individuals, including clinicians, were presented with outcome data for a variety of treatments on offer. The situation was hypothetical with the aim of examining choice preferences and inconsistencies by any

Figure 16 **Relationship between judged frequency and the actual number of deaths per year for 41 causes of death**



Source: Lichtenstein *et al* (1978). *Note:* Plotted on logarithmic scale.

of the groups when selecting treatment options.

The most notable finding of the study was that when two otherwise identical treatments were presented, but treatment A was described as having a one year *survival rate* of 90 per cent and treatment B of having 10 per cent *mortality* at one year, then all groups of respondents – including clinicians – expressed a consistent preference for 90 per cent survival rather than 10 per cent mortality. The expected outcome of the two was the same, only the terminology was different.

The authors could offer no explanation of this interesting phenomena and comment that '. . . this effect of using different terminology to describe outcome represents a cognitive illusion'. The fact that both doctors and patients may suffer from such a cognitive illusion suggests that this area is worthy of further study. If the 'framing' of treatment problems in terms of probability of living or probability of dying can systematically bias perceptions and preferences then it seems likely that the quality of medical decision making would benefit from more detailed investigation of this cognitive illusion.

This question of the 'framing' of outcomes influencing perceptions and preferences was investigated by Eraker and Sox (1981) who examined the preferences of individuals in hypothetical pharmaceutical decisions where treatment outcomes were uncertain. In each treatment 'scenario' individuals were asked to choose between two medicines. Treatment outcomes were expressed in terms of life-expectancy; medicine 'C' offering a 'certain' outcome and medicine 'U' offering an uncertain outcome – essentially a gamble between high and low life-expectancy. Overall the expected outcome of medicine C was the same as medicine U.

The results of various trade-offs between the certain (C) and the uncertain (U) medicines are presented in Table 14. In four cases out of five the certain effect has the same expected outcome as the uncertain 'gamble' but respondents consistently chose the certain option. This 'certainty effect' is therefore a demonstration of risk aversion. Individuals prefer not to gamble even if the expected outcome is the same. As discussed earlier there is a *process* utility at work here which cannot be divorced from the outcome utility.

Note however, that in this first experiment Eraker and Sox have expressed all treatment outcomes in terms of probabilistic *therapeutic* effects from therapy, such as

Table 14 **Patient preferences for palliative effects of medication for serious chronic illness**

Comparison of expected values of certain (C) and uncertain (U) choices	*Certain therapeutic drug effect*	*Uncertain therapeutic drug effect*			*Total number of patients responding*	*Proportion of patients selecting certain drug effect*
C = U	1 extra year of life	0.67 chance at 1.5 extra years of life	AND	0.33 chance of no extension of life	382	0.73[a]
C = U	1 extra year of life	0.50 chance at 2 extra years of life	AND	0.50 chance of no extension of life	378	0.61[a]
C > U	1 extra year of life	0.50 chance at 1 extra year of life	AND	0.50 chance of no extension of life	378	0.94[a]
C = U	2 extra years of life	0.50 chance at 1 extra year of life	AND	0.50 chance at 3 extra years of life	375	0.68[a]
C = U	2 extra years of life	0.50 chance at 4 extra years of life	AND	0.50 chance of no extension of life	376	0.69[a]

[a] $p > 0.001$ that observed proportion differs from 0.50 by chance.

Source: Eraker and Sox (1981)

gains in the life expectancy. Would changing the way in which the treatment gambles were 'framed', moving from 'gain' terminology to 'loss' terminology, influence the respondents' attitudes to risk?

The pioneering work of Kahneman and Tversky (1980) on **prospect theory** as an alternative to expected utility theory examined attitudes to risk when gambles were expressed using gain and loss terminology. They found that individuals were *risk averse* when gambles were expressed in terms of *gains* but *risk lovers* when they were expressed in terms of *losses*. Expected outcomes may therefore be the same for particular gambles with attitudes to risk being determined by the terminology of the choice. The distinction is similar to the old adage about the optimist who states that the glass is half full and the pessimist who says it is half empty. Depending upon which side of the coin is emphasised, choices involving risk will vary.

To return to the pharmaceutical gambles, Eraker and Sox now examined how individuals would respond in two different scenarios. In the first scenario, individuals must again choose between a 'certain' treatment outcome and a gamble where the gamble outcomes are expressed in terms of probabilities and *therapeutic effects* (gains). In the second scenario, the individual must choose between the 'certain' outcome but this time the alternative is a gamble where the pay-offs are presented in terms of probabilities and *adverse effects* from the medicine. For example, an individual might have to take a decision about whether or not to take a treatment for a headache. The treatment will reduce the duration of the headache with probability (P) and lengthen the duration with probability (1–P). The findings are that individuals will tend to be risk *averse* if the situation is framed as a gain (*reduction* of duration of headache) and as risk lover if it is framed as a loss (*increase* in duration of headache).

The results from the Eraker and Sox (1981) experiment are presented in Table 15. In all cases where the

Table 15 **Patients' preferences for therapeutic and adverse drug effects**

Scenario	*Mean number respondents to each question in a scenario (range)*	*Certainty preference score (mean score/patient ± SD)*[a]
Therapeutic drug effects		
Drug prolongs life in fatal illness (no side effects)	378 (375–382)	0.66 ± 0.35[b]
Drug reduces duration of severe headache (no side effects)	459 (454–467)	0.66 ± 0.35[b]
Drug improves exercise tolerance by preventing chest pain (no side effects)	436 (380–457)	0.67 ± 0.36[b]
Adverse drug effects		
Reduction in life expectancy due to effects of drug treatment for fatal illness	340 (335–350)	0.37 ± 0.35[b]
Drug that relieves headache causes nausea and repeated vomiting	432 (425–441)	0.39 ± 0.31[b]
Drug that relieves angina reduces exercise tolerance by causing exertional dyspnea	394 (343–415)	0.44 ± 0.34[b]

[a] A patient's certainty preference score is the proportion of the four questions in which the patient chose the 'certain' alternative.

[b] $P < 0.001$ that the observed proportion differs by chance alone from 0.5 (proportion to be expected if patients were indifferent between the 'certain' and 'uncertain' alternatives).

Source: Eraker and Sox (1981)

alternative gamble to the certain equivalent was expressed in terms of therapeutic (gain) terminology, respondents consistently preferred not to gamble and were risk averse. Where the terminology was switched to that of *loss* and adverse effects that might arise, the respondents consistently preferred the uncertain option of the treatment gamble.

A similar experiment was conducted by Breyer and Fuchs (1982) who asked 325 subjects to select one of two treatments for a hypothetical illness. One treatment had a certain outcome, the other was a gamble. Outcomes were expressed as episodes of pain or pain relief. Again they found that individuals were risk averse with respect to gains and risk loving with respect to losses. These findings being consistent with Eraker and Sox, and with the experiments that Kahneman and Tversky (1982) have undertaken with choice experiments involving money.

Although this sort of bias in decision making under risk can be observed, it is difficult to explain. Indeed, it goes well beyond the terrain of economics and into psychology and the study of cognition. The problem of risk perception, or more precisely of mis-perception, is therefore not only due to the way in which people mis-judge likelihood, (which may in turn be due to variable relationship between numeric and verbal communication of probability) but is also due to the fact that the way in which outcomes are framed – gains or losses. This area is worthy of further multi-disciplinary research.

VOLUNTARY OR INVOLUNTARY RISK: A QUESTION OF CONTROL

Another important element in the evaluation of different risks is the question of whether risks are perceived as being *voluntary* or *involuntary*. The difference between the two types of risk being a function of *controllability*; that is, the person's ability (as he or she sees it) to influence or control the particular risk.

Voluntary risk examples include smoking, rock climbing and parachuting; all of which are activities that the individual might *choose* to do, having weighed up the expected benefits and costs (including injury and fatality risk) to him or her. Such risks are voluntarily accepted by the individual and not forced upon them. Involuntary risks are those which we have little or no direct control over at the individual level, the 'acceptability' of risk levels being determined collectively. Examples of involuntary risks include natural disasters (fires, floods) and risks of leaks from nuclear power plants.

Chauncey Starr (1969) examined the nature of voluntary and involuntary risks and argued that in order to attain some given level of benefit, the public seems willing to accept risks from voluntary activities which are roughly a thousand times greater than it would from involuntary activities. Therefore the extent to which individuals perceive a risk as being voluntary will influence the degree of benefit-risk trade-off they are prepared to accept.

An interesting question, pertinent to risk/benefit trade-offs in medicine, is whether health care risks are viewed as voluntary or involuntary. On one level it is possible to view all medicine as voluntary risk taking by the patient. The individual seeks medical advice and diagnosis, consents to, and complies with treatment. The element of control is that the patient can decide not to consent to treatment if he or she feels that the risks are too high relative to their perception of benefit. But this obviously raises the question about the amount of risk information the doctor or surgeon should provide to the patient.

Does 'informed' consent mean that *all* risks should be disclosed to the patient prior to the treatment? If not, *how much* risk information is the doctor obliged to disclose to the patient?

8 TREATMENT RISKS AND INFORMED PATIENT CONSENT

The National Consumer Council (1983) in its volume entitled **Patient's Rights** gives guidance to patients and doctors concerning consent for operations:

> '. . . Doctors must tell patients any facts necessary for them to decide whether they want the operation. Exactly how much the doctors tell is a matter of discretion, but real and foreseeable risks should be disclosed.'

More generally the NCC go on to offer guidance on what information should be made available to patients for any treatment:

> 'The general rule is that if patients ask for information, their questions should be answered fully and truthfully. But the amount of information doctors give patients varies according to their assessment of how much a particular patient wants to know, and how much they think the patient can handle. Doctors' duty of care towards patients includes volunteering information of the real risks of particular treatments, which patients need in order to give their consent to that treatment.'

The medical profession faces something of a dilemma where disclosure of treatment risks is concerned. When the medical practitioner offers treatment he will have a body of technical information on the likelihood of different outcomes. How much of this probabilistic data should he spontaneously proffer to the treatment candidate before consent can be said to be 'informed'? Should patients be told of *all* treatment risks or should doctors take the responsibility for judging which risks are significant and therefore relevant for the patient to consider when deciding whether to accept treatment?

The aim of this section is to review the question of informed patient consent in the context of information about treatment risks. Again the focus is on two applications of medicine: prescription medicines and surgery. With pharmaceuticals there is increasing pressure on manufacturers to enclose package inserts which contain detailed risk information for patients. How much information should be supplied, and will the availability of such data *educate* the consumer about risks or *frighten* the patient into non-compliance?

The surgical discussion focuses on the recent case that went to Appeal and finally the House of Lords (Sidaway *v* Bethlem Royal Hospital). The opinions given by the Lords serve as an illustration of the legal and medical profession wrestling with the definitions of risk as they apply to clinical practice. An interesting question emerges for debate: should the UK adopt a US-style informed consent doctrine with the attendent prospect of litigation and malpractice suits, or continue with the present system where the last word remains medical not legal, disclosure of risk information being held as a 'clinical judgement'?

PHARMACEUTICALS AND PACKAGE INSERTS: WILL THEY EDUCATE OR FRIGHTEN?

In the interests of promoting 'safe and effective' use of oral contraceptives the US pharmaceutical regulatory body, the Food and Drugs Administration (FDA), has required that information on the risks of such preparations be included in the form of patient package inserts (PPIs). These inserts are written in non-technical language to inform the user about potential risks and benefits of the medicine. The motivating principle behind the move is that;

> '. . . patients have a right to know about the effects, positive and adverse, of prescription medicines, and that such information will promote the safe and effective use of such products.' (Keown *et al*, 1984.)

The authors also note that in 1979 the FDA pledged to expand this patient information programme to 'most prescription medicine products for human use'.

Surveys of consumers have tried to determine the demand for such information. Joubert and Lasagna, (1975) for example, found that most respondents (93 per cent) wanted to know the reasons for the medicines they were using, the common risks involved (89 per cent), the risks of under and over-dosage (82 per cent) and the probability of rare side effects occuring (81 per cent).

The question of supplying risk information on PPIs is therefore not one of all or nothing, but of how much? The Joubert and Lasagna study found that 81 per cent of patients surveyed wanted to be informed about fatality risks even if they were as low as 1 in 100,000. Mazis *et al* (1978) surveyed oral contraceptive users and found a preference for longer and more detailed PPIs. The indication from such surveys is that the public demand for pharmaceutical risk information is larger than that perceived (or believed appropriate) by the prescribers and suppliers of medicines. Keown *et al* (1981) found that doctors and pharmacists preferred a policy of disclosing only serious side effects and common minor effects on the PPI whereas the lay person preferred to be informed of *all* known side effects, regardless of probability and severity.

On the question of how much information PPIs should contain, the FDA proposed that only serious and frequently occuring adverse effects would be listed. (US Federal Register, 1979). This judgement, to keep PPIs short and simple, reflects concern amongst policy makers that although patients demand a high degree of risk information the general public's ability to digest large quantities of such data is extremely limited.

Patient compliance with pharmaceutical regimes will be influenced by an individual's *perception* of the risks attached to a particular medicine. As discussed in the previous section, individuals introduce a wide range of heuristics and biases when making judgements concerning probabilities. Of particular relevance here is the general inability to comprehend very small probabilities. Often the probable and the possible can become blurred. Detailed information presented in a package insert on what *might* happen could frighten more patients than it educates.

But should a generalisation such that 'people do not understand risk data' put an end to attempts to inform

the public about risks using PPIs? More specifically is there any evidence that greater information in pharmaceutical packages is associated with increased patient non-compliance?

The evidence to date is largely anecdotal. Guarino (1979), for example, when discussing PPIs for hypertensive patients, comments that, 'Clearly we just don't know yet whether patient compliance will be improved as a result of this type of information', and he goes on to call for '. . . careful, deliberate and scientific study of these problems'. Further research in this area would clearly be valuable. There is a need to expand our knowledge of public perceptions of medicine risks with respect to different methods of presenting risk data such as package inserts.

Keown *et al* (1984) have provided an initial analysis of risk perception and PPIs. They attempted to test the hypothesis that 'the presentation of a lengthy list of rare side effects will cause people to see a medicine as riskier than their physician sees it'. Data on the frequency and type of side effects for six hypothetical medicines were presented to a group of physicians and lay people; they were asked to rate the medicines on the basis of perceived risk. The results of this small study indicated that there were marked discrepancies between the perceptions of the lay-people and the doctors when presented with the same information. In particular, and as predicted, there was evidence that the public tend to attach relatively more weight to rare events than do doctors.

But the observation that doctors and patients differ in their perceptions of risk does not, of itself, point to the prescription that treatment choices should ignore the patient's attitude towards risks in treatment decisions. This topic is further discussed in the section on risk evaluation below.

THE DOCTOR IN THE MIDDLE

In the absence of PPIs it is the prescribing doctor who informs the patient about the risks associated with a medicine. The doctor judges how much information to give the particular patient. He or she is in the middle of a triangle between the manufacturer, the government (CSM) and the patient.

If the doctor does not warn the patient of a risk and it materialises, where does the responsibility for risk lie? Should the patient seek legal redress from the doctor, the government or the manufacturer?

In practice it is typically the last, with the responsibility for risk lying with the industry. This seems reasonable if the product can be shown to be in some way *defective*. (If you buy a television and it doesn't work then you expect your money back). But if the medicine was not defective and it just happens that the victim was the unlucky patient in a million where the risk materialises, what then? Who is responsible for this gamble in which the victim lost? The Government (Medicines Division) has licensed the medicine for use, the CSM advised them and they judged the risk level as being acceptable; might not some responsibility lie with them?* This question of risk responsibility in medicine is further discussed below.

SURGERY AND MATERIAL RISK: THE CASE OF SIDAWAY

The vast majority of surgical operations are elective and patients' have the choice of refusing consent. When faced with the decision to sign the pre-operative consent form the main source of information available to the patient is the surgeon. How much information should the surgeon provide, and can a surgeon actually withhold risk information from a patient if he thinks it in the best interest of the patient? These points and other illuminating medico-legal aspects of risk were recently given an airing in the House of Lords when their Lordships gave their opinions on the case of Sidaway *v* Bethlam Royal Hospital.

In 1974 Mrs Amy Sidaway had an operation on her cervical spine in order to relieve acute neck and shoulder pain. The operation carried a risk of one or two per cent of causing damage to the spinal column and nerve roots. The operation was not performed negligently by the neurosurgeon Mr Murray Falmer but the risk materialised and left the patient paralysed. She sued the hospital and the surgeon on the grounds that she had not been adequately warned of the risks prior to the operation and therefore her consent to the operation was not 'informed'.

Unfortunately the surgeon died in 1977 before the trial. The judge rejected Sidaway's claim that no mention of the risks had been made. The evidence was that the surgeon had explained the possibility of nerve root damage but had not mentioned the more remote danger of spinal cord damage and paralysis. The amount of explanation given was judged to be in accordance with 'accepted medical practice' and on the **Bolam** test (Bolam *v* Friern Hospital Management Committee [1957] 2 A11 ER 118) the doctor could not be held to be negligent. This **Bolam** test was accepted in the High Court and Mrs Sidaway's financial loss was estimated at £67,500.

Sidaway then took her case to the Court of Appeal who also endorsed the **Bolam** test: disclosure of information to patients was a matter for clinical judgement. Finally Mrs Sidaway went on to the Law Lords, where her case was dismissed, but not without producing some interesting opinions from their Lordships concerning the relationship between legal and medical definitions of informed consent and risk. Lord Scarman was the one Lord who dissented from the general view that the disclosure of risk information to patients is a matter of clinical judgement.

The views of the Law Lords on this topic of informed consent are fascinating. However, because these legal details can be skipped by the general reader at no great loss, the issues have been summarised in the box on page 37: Risk, the Law and Informed Patient Consent.

SUMMING UP

It is difficult to determine where the balance lies on the issue of how much risk information to give patients – either for medicines or surgery. At the heart of the problem is the nature of the doctor-patient relationship. In economic jargon the traditional view is that the doctor acts as the patient's agent and uses his skills to maximise the patient's expected utility or wellbeing. The doctor behaves in an 'optimal' fashion inasmuch as his behaviour would accord with that of the 'perfectly prudent patient' – if such a thing existed.

If such a relationship is to function efficiently then the doctor must be aware of the patient's preferences on a whole range of issues – one of which being his or her attitude to risk. The problem at present would seem to be how best to communicate risk information to patients

*In fact, at the time of writing, a thousand people damaged by the withdrawn anti-arthiritic compound Opren are taking legal action for compensation against both the manufacturer *and* the government. For a recent discussion of pharmaceutical liability issues, see Newdick (1983).

RISK, THE LAW AND INFORMED PATIENT CONSENT

In the case of Mrs Amy Sidaway the first issue to tackle was whether or not the surgeon had been negligent – could his behaviour be construed as malpractice? The expert opinion was that the operation had not been performed negligently and the patient had just been the unlucky case where the risk materialised. The question then becomes one of how widely malpractice is defined and whether this includes failure to inform patients of risks.

The majority view of the Lords was that the amount of risk information given to a patient should be a matter for clinical judgement. The one dissenting view was that of Lord Scarman who argued in favour of an 'objective test' similar to the informed consent doctrine adopted by some States in the USA (Canterbury *v* Spence, 1972). Such a test would be objective inasmuch as it would be independent from clinical judgement.

Lord Scarman posed two fundamental question:

1 'Has the patient a legal right to know, and is the doctor under a legal duty to disclose, the risks inherent in the treatment which the doctor recommends?'

2 'If the law recognises the right and obligation, is it a right to full disclosure or has the doctor a discretion as to the nature and extent of his disclosure?'

The view expressed by the *Bolam* test clearly gave no proviso for a legal right to the patient for such information and this is the source of Scarman's criticism:

> 'It leaves the determination of a legal duty to the judgement of doctors. Responsible medical judgement may, indeed, provide the law with an acceptable standard in determining whether a doctor in diagnosis or treatment has complied with his duty. But is it right that medical judgement should determine whether there exists a duty to warn of risk and its scope?'

Scarman therefore considers that doctors have a legal duty to disclose information about risks to patients. But how much information? Scarman goes on to rehearse the propositions laid down in *Canterbury v Spence* (USA), an important clause being that the doctor must, therefore, disclose all 'material risks'. What risks are 'material' is determined by the 'prudent patient' test which was formulated as:

> 'a risk is . . . material when *a reasonable person*, in what the physician knows or should know to be the patient's position, would be likely to attach significance to the risk or cluster of risks in deciding whether or not to forego the proposed therapy.'

Superficially (and, perhaps, uncharitably) this appears to be replacing the medical autonomy of the *Bolam* test with legalistic vaguery. With differing perceptions and attitudes to risks, based on different experiences, it is difficult to grasp quite who the 'reasonable person' of the *Canterbury v Spence* test would be. Assuming, for the moment that such a test can be applied and a surgeon is found 'guilty' of not informing a patient about a 'material' risk prior to an operation, is the doctor liable? No. The final proposition of the test is what the court called a 'therapeutic privilege'. In summary:

> 'This exception enables a doctor to withhold from his patient information as to risk if it can be shown that a reasonable medical assessment of the patient would have indicated to the doctor that disclosure would have posed a serious threat of psychological detriment to the patient.'

So even if it can be shown that the risk is 'material', if the knowledge of this risk is likely to be detrimental to the patient's health the doctor can withhold it. This remains the legal viewpoint and has been put even more clearly by Lord Denning in his book 'The Discipline of Law' (1979) where he discussed the case of a female broadcaster who underwent surgery on her thyroid gland. Before the operation she was reassured about the risks but in the event a nerve was so badly damaged that she could no longer speak properly. Lord Denning poses the question:

> 'What should the doctor tell his patient? Mr Tuckwell admitted that on the evening before the operation he told the plaintiff that there was no risk to her voice when he knew that there was some slight risk, but that he did it for her own good because it was of vital importance that she should not worry. In short, he told a lie, but he did it because he thought in the circumstances it was justifiable.' (p 243)

The dilemma is not legal or medical but philosophical; is it morally sound that doctors should not tell the truth because they believe it is in the interest of the patient? There are arguments for and against this sort of 'benevolent deception'. The fact remains however, that in the eyes of UK law the disclosure of risk information remains a clinical judgement.

[Interested readers should refer to the excellent book on paternalism in health care by Childress (1982).]

without causing undue panic and concern that might unnecessarily threaten compliance. The doctor-patient relationship remains one where the doctor assumes complete responsibility – including the judgement about how much risk information to give patients.

Therefore, in response to the question raised in the previous section, treatment risks are *voluntary* only to the extent that the individual doctor sees fit to disclose risk information to the patient. UK law holds this to be a matter for clinical judgement alone.

The alternative is to alter the doctor-patient relationship, sharpening its contractual premise. But the likely cost of introducing a US-style informed consent doctrine is that the floodgates of litigation would be opened. Medical malpractice suits in the area of informed consent are big business in the US. With growing claims against doctors, the tendency is for medicine to become far more defensive in fear of litigation. The last word therefore rests with Lord Denning (1979):

> 'It is, I believe, very different in the United States of America. "Medical malpractice" suits there have become the curse of the medical profession. The legal profession get "contingency fees". So they take up cases on speculation. The jury gives enormous damages. Insurance premiums are high. The doctors charge large fees to cover them. It is all very worrying.'

EVALUATING HEALTH CARE RISKS – A QUESTION OF RISK BALANCING

All medicine is a calculated risk. The weighing-up of the risks of treatment against the risks of no intervention: a question of risk balancing. The risks that individuals are prepared to accept will be in proportion to the expected benefits gained in terms of improvements in life-expectancy and/or quality of life. But what factors do or should influence the extent to which patients and doctors are prepared to enter such lotteries?

The aim of this section is briefly to review the ways in which risks are evaluated in medicine. Then to examine the role of patient preferences in the balancing of treatment risks. Particular attention is paid to the question of doctor and patient *attitudes* to risk when comparing expected treatment outcomes.

ACCEPTABLE OR ACCEPTED RISKS?

Various risk commentators have reviewed the ways in which risks can be evaluated (see for example Jones and Akehurst, 1980). The trade-off between risks and benefits can be discussed at the macro (collective) level or the micro (individual) level. An example at the macro level might be to determine what frequency and type of adverse drug reactions are 'acceptable' before a medicine is licensed for marketing. The risk regulators need to form some judgement, on our behalf, concerning the 'acceptability' of particular risks.

At the micro level the choice between different treatment regimes also involves a judgement about relative risks. As demonstrated earlier in this paper, treatment choices can be analysed in terms of expected pay-offs in units of quantity and quality of life. Depending upon the extent to which the patient's preferences regarding outcomes and the process of risk taking are used, the judgement of the acceptability of the risks associated with a given treatment will be more or less a judgement made by the clinician.

An important question to be addressed therefore, at both the macro and the micro level, is *how* judgements about the acceptability of risks are made. Essentially this is a question about how, at the collective and individual levels, risks are evaluated.

One approach for judging the acceptability of a particular risk is to compare that risk with those that people *already* run and presumably accept. The logic of this line of reasoning follows from the concept of **revealed preference** in economics, where observed choice patterns indicate or reveal the underlying preferences that individuals hold. If the risks associated with new medicine A are fewer than those associated with existing medicines B to Z which are already accepted and offer similar therapeutic benefit, then on this simple premise the risks of medicine A can be judged to be 'acceptable'.

Kletz (1977) has termed this principle that of the 'accustomised' or accustomed risk, where risk policy guidelines are computed as threshold values – levels of risk to which society has become accustomed.

A number of authors have attempted to determine which levels of risk society has become accustomed to, and therefore might be judged as being acceptable. Rothschild (1978) noted that we accept a fatality probability of 1 in 7,500 per year from motor accidents and puts this forward as an acceptable level. Therefore, any risk falling below this should be accepted if society is being consistent in terms of risk preferences. Similarly, Knox (1975), examined a number of risks including medicial risks, and stated that:

> 'The general conclusion from these examples together, and especially the medical examples, is that a risk of about 10^{-5} consequent upon a single decision is somewhere near the level below which concern ceases; that is the level at which in common parlance, a procedure is regarded as safe.'

Here Knox is advocating that any fatality risk which has a frequency of less than 1 in 100,000 is perceived as being 'safe'. Below this level 'concern' ceases. This seems rather a hurried 'rule of thumb' judgement which ignores the nature of different risks: some are voluntary, others involuntary. As Starr (1969) has demonstrated, 'concern' about risks is not simply a function of likelihood but is also a function of the degree of control an individual has over risks.

There are a number of weaknesses in the reasoning behind the acceptable risk approach. Fischoff *et al* (1979) note that the approach 'assumes that past behaviour is a valid predictor of present preferences, perhaps a dubious assumption in a world where values can change quite rapidly'. Furthermore, this revealed preference method is based on the assumption that people have full information about the alternatives that are available. This is unlikely to be the case. Therefore it may be wrong to argue that because particular risks are *accepted* by society, this is then an indication that these risks are *acceptable*.

Jones and Akehurst (1980) have also levelled criticisms against this threshold approach of acceptable risk levels. Their main criticism is that such an approach concentrates exclusively on the cost (risk) side without reference to the benefits. Policies and decisions which use this one-sided approach are likely to be inefficient because no account is taken of the relationship between expected benefit and expected cost.

Under the umbrella of the cost-benefit approach as a way of thinking about evaluation there exist specific techniques such as cost-benefit analysis (CBA), and cost-effectiveness analysis (CEA). A subsidiary technique often cited in the risk literature is that of risk-benefit analysis (RBA), this being a special case where some of the costs of a decision or option have probabilities attached to them and are presented as risks. As with CBA, the use of RBA would be to determine those options or choices which provide the greatest surplus of benefit over risk. Comparing options will illuminate the ways in which individuals are prepared to trade-off benefit against risk at the margin.

An important facet of the cost-benefit approach as a means of evaluating risk-benefit problems is that values and preferences must be explicitly attached to outcomes. In addition, for the calculation of expected utility, the individual's preferences regarding the *process* of risk taking must be included. But as Fischoff *et al* (1981) have noted, '. . . risk attitudes have little place in the theory or practice of cost-benefit analysis'. This shortcoming has led to greater use of *decision analysis* in the evaluation of clinical practice; a technique which draws on aspects of economics, operational research and management science.

Evaluation problems in clinical practice are being increasingly addressed using decision analysis (see Weinstein *et al*, 1980), where treatment strategies can be mapped out as a series of probabilities and outcomes on tree diagrams to calculate expected outcomes and values. Furthermore this mode of study has helped to illuminate the whole area of the relationship between patient preferences regarding risk taking, therefore enabling clinicians and researchers to use information on patient attitudes to risk in order to calculate expected utilities for treatment strategies.

The next section offers examples which illustrate the importance of the way in which treatment outcomes are presented to patients and also the way in which patients' attitudes to risk can influence treatment choices.

EVALUATING SURGICAL RISKS: FALLACY OF THE FIVE-YEAR SURVIVAL RATE

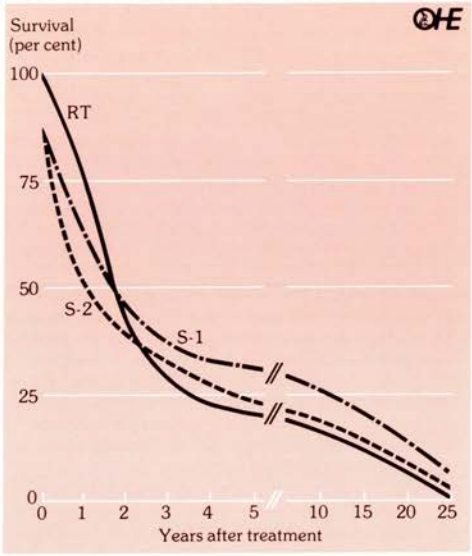
In a series of articles Barbara McNeil and colleagues from Harvard University have explored the way in which patients' attitudes to risk and other preferences might be incorporated into treatment choices. The subject of one such study was the choice between radiotherapy and surgery for treatment of lung cancer – McNeil *et al* (1978). The 'best' outcome that a patient could expect with surgery was a five-year survival rate of 33 per cent and this compares with the five-year survival prospects with radiotherapy of 21 per cent.

If the choice criterion is longevity and a patient is presented with these data only, then the presumed choice of treatment would be surgery. However, merely quoting the five-year survival rate ignores the *distributions* of survival which are associated with the two treatment modalities and hence the distribution of risks that a patient will be exposed to. The surgical option offers better *long-term* survival prospects but there is an immediate operative fatality risk of 10 per cent which patients are exposed to. The surgery option is a gamble which is expected to pay off but the question of whether this gamble is acceptable is dependent upon the patient's (or doctor's?) attitude to risk.

The survival comparisons for surgery and radiotherapy are presented in Figure 17 with the survival curves for the two modalities crossing-over at around two years after treatment. The attitude to risk that McNeil *et al* found among patients with operable lung cancer was that of considerable risk aversion: patients preferred not to gamble with the risk of operative mortality even if this would mean lower survival in the long run.

This is a finding which runs counter to the surgeon's intuition and inclination to prefer those treatments which offer the greatest life-expectancy. The distribution of risks and the patient's attitude toward such risks in not generally a consideration. Yet to ignore patient preferences on risk might be to include people in treatment gambles that they may prefer to avoid. The opposite case will also apply. A patient may be prepared to gamble on treatment whereas the clinician may be risk averse.

Figure 17 **Survival curves for operation and radiation therapy for 60 year old patients**



On the ordinate is the cumulative percentage of patients surviving, and on the abscissa is the time from treatment. Curve S-1 represents excellent surgical results, and curve S-2 typical results; both surgical curves include a 10 per cent operative mortality rate. Curve RT summarises the survival after radiation therapy. Note the marked difference in the shape of the survival curves for operation and radiation therapy during the first two years after diagnosis.

Source: McNeil *et al* (1978).

A further reason why life-expectancy might not be the best measuring rod for evaluating treatment decision problems is that outcomes will differ in terms of quality of life. McNeil *et al* (1981) have explored what preferences people would have in the treatment of laryngeal cancer. Again, the choice is between surgery (laryngectomy) with a three-year survival rate of 60 per cent and radiation therapy with a survival rate of 30 to 40 per cent at three-years. The other outcome difference however, is in quality of life. The laryngectomy leads to a loss of normal speech while radiation therapy leaves the voice intact. Given this trade-off problem between quantity and quality of life, the researchers found that a fifth of the healthy volunteers interviewed would choose the radiation therapy in preference to surgery. Maximising *health-expectancy* is not necessarily synonymous with maximising *life-expectancy*.

Obviously not all treatment decisions are as clear-cut as these examples. In many instances there may be effectively no alternative treatment strategy (other than doing nothing) with which to make comparison. Treatment A may *dominate* treatment B in terms of risks *and* benefits, in the view of both the clinician and the patient.

But where treatment choices do exist, with differential risks and benefits, the choice of one treatment in preference to another is to invoke *somebody's* attitude to risk in relation to benefit. Values are being attached to outcomes

and risk-taking *implicitly* by the clinician if he makes the decision. It may be the case that the clinician's preferences – his desire to gamble, how he rates quality of life relative to quantity of life – are identical to those of his patient. But this is unlikely. Only by eliciting such preferences from patients can *explicit* values be placed on these aspects of choice as a means of improving medical decision making.

DOES THE DOCTOR KNOW BEST?

Such ideas go against the traditional grain of medicine where the assumption has always been that the 'doctor knows best' for his patient. The new wave of clinical decision analysis – of which the Harvard School is a leading example – is bringing the utilitarian principles which underpin economics into the sphere of the doctor-patient relationship. One of the guiding principles of economics is that individuals are assumed to be the best judges of their own welfare – people know what is good for them. This idea of 'consumer sovereignty' is not a central theme in the traditional paternalistic model of medicine (Childress 1982).

This is not to suggest that the patient will have any detailed knowledge of modern medicine – such expertise is the doctor's. Rather it is to view the doctor as a producer of 'health' – an entity which can be produced in many different ways with many different characteristics. The question is, should the characteristics of the product and chosen method of production with their varying degrees of risk reflect the preferences of the consumer or the producer?

Many would agree to patient preferences on risk in principle but disagree in practice, arguing that the doctor's understanding of statistical and probabilistic data will be superior to that of the patients, and therefore he or she should judge. Yet the example presented earlier – where clinicians preferred a treatment with 90 per cent survival in favour of a treatment with 10 per cent mortality – McNeil *et al* (1982) – suggests that the medical profession are not immune from bias and mis-perception of risks.

Obviously the medical practitioner has an armoury of technical knowledge concerning outcomes and their likelihood. The patient relies on the doctor's expertise for information. But where treatment choices exist it would appear that the quality of clinical decision making might be improved if patient preferences were elicited. In some instances this may already be the norm but the evidence suggests it is the exception.

SHIFTING RISK RESPONSIBILITY: AN EXAMPLE

It is interesting to speculate on the likely outcome of shifting the responsibility for risk-taking in medicine more towards the individual patient and away from the doctor or some third-party risk regulator. Such a consideration, essentially allowing individuals more freedom to gamble on treatments, is the logical extension of including patient preferences in treatment decisions.

The idea can best be illustrated by way of an example. The recent controversy surrounding the risks of Opren serves as a prime example of a medicine which produced substantial health benefits to arthritis sufferers being withdrawn from the market due to its risks.

The medicine was withdrawn by the company after the CSM advised the government temporarily to suspend the product license when the medicine had been associated with a number of ADRs, some fatal. There is not space enough here to debate how accurately the risks of Opren had been assessed, although the weaknesses of voluntary reporting mechanisms have already been mentioned. Inman (1985) for example, when examining the extent to which the reported fatal ADRs could be attributed to the medicine has argued that in reality the risks were quite small. In particular, the risks were largely limited to elderly patients for whom the recommended daily doses of 600mg were effectively over-doses.

But what of the benefits of Opren? The large marketing success of the medicine prior to withdrawal was due to the fact that it rapidly became accepted as an effective means of relieving the pain of arthritis. The costs of withdrawing the medicine can be seen in terms of the health benefits that such patients must forego. As Teeling Smith (1982) commented:

> 'Certainly its withdrawal has prompted a great deal of anecdotal evidence of patients who were prepared to accept its adverse effects for the sake of its exceptional benefits. Two "satisfied patients", for example, subsequently wrote to *The Times* arguing that the relief obtained was so great that patients should still be able to take the medicine, and to accept individual responsibility for the risks involved.'

The indication is that some individuals would be prepared to accept a shift in the responsibility for risk.

In many circumstances people are prepared to gamble on the risks of mortality in order to gain some improvement in quality of life. But this is not surprising. Such trade-offs are made every day of our lives in areas outside of health care. The man that crosses the busy road without walking down to the zebra-crossing is trading-off convenience and time against the risk of death or injury. But that's a choice which reflects *his* preferences and perception of the risks. Could not more freedom of choice be introduced into health care?

It would be naive to assume such changes could occur overnight. The preliminary research and education agenda is obviously large. More needs to be known about how individuals (doctors and patients) comprehend and perceive risks. This knowledge can then be used to devise methods for communicating risk data more effectively. Only then can the responsibility for treatment risks be moved more gently towards the patient.

10 DISCUSSION: PREDICTING THE FUTURE

As medicine and health care move towards the twenty-first century, treatment would appear to be increasingly less about extending the length of life and more about improving its quality. But medicine is not immune from the Law of Diminishing Returns. It seems reasonable to argue that the future incremental gains in health – reductions in pain and disability – will be purchased only at the price of increased risk-taking. As with all other human endeavours, progress depends upon taking chances.

This is not to suggest that the existing catalogue of treatments are becoming more risky. As has been demonstrated the trends over time are downwards. Rather it is to recognise that new innovations, be they surgical or pharmaceutical, often carry substantial new risks. More than ever before this forces the issue of *how* costs, risks and benefits should be evaluated and also *who* should be responsible for making such trade-offs.

This review of clinical risks has raised more questions than it has answered. Given the paucity of research in this area – especially in the UK – this is not surprising. There would appear to be an enormous rift between the professional assessment of risk and the public's perception and understanding of the concept. This is cause for concern for the medical practitioner because it is he or she who faces the task of translating the technical into the non-technical for each patient.

This task is made more difficult by the fact that statistics hold little meaning for most people. On the other hand, verbal representations of risk and probability are heavily value laden: 'likely' will be interpreted differently by different people. There is no simple 'scale language' of risk.

Perceptions of risk will also be dependent upon how treatment choices are 'framed'. To choose a treatment which has 90 per cent *survival* in preference to one which has 10 per cent *mortality* (other things being equal) can be described as a 'cognitive illusion' – McNeil *et al* (1982) – but it is worrying that such a bias can be consistently implemented by doctors and patients alike.

Yet although our understanding of risk is often poor, the demand for more risk information seems ever increasing. The question is not one of all or nothing but of *how much* risk information to give patients so as to educate rather than frighten. Again, this is an area for more research. The question of whether pharmaceutical package inserts of different types will affect patient compliance is largely an empirical one.

A number of the studies reviewed in this paper have addressed the important issue of incorporating patients' attitudes to risk into treatment decisions. This is an area of preference which is often overlooked. It is complacent to assume that a patient's desire to gamble on a treatment will be the same as the clinicians. The McNeil studies show that patients are often averse to taking risks and are prepared to trade longevity for quality of life improvements.

But, if the long-term objective is one of incorporating patient preferences more into clinical decision making, then the obvious pre-requisite is that doctors and (potential) patients should be made more aware of the nature of risks as they relate to the benefits of treatments. To this end, the research and education agenda is large but the expected pay off will be in terms of improvements in the quality of clinical decision making.

REFERENCES

Beyth-Marom R (1982)
'How probable is probable? A Numerical Translation of Verbal Probability Expressions', *Journal of Forecasting*, vol 1, p 257–269.

Bolam *v* Friern Hospital Management Committee (1951)
2 A11 ER 1118.

Breyer R and Fuchs V (1982)
'Risk Attitudes in Health: An Exploratory Study', National Bureau of Economic Research, Working Paper No 875, Cambridge, Mass.

Buxton M, Acheson R, Caine N, Gibson S and O'Brien B (1985)
Costs and Benefits of the Heart Transplant Programmes at Harefield and Papworth Hospitals, HMSO.

Canterbury *v* Spence (1977)
464F 2d 772.

Childress J F (1982)
Who Should Decide? Paternalism in Health Care, Oxford University Press.

Cohen A V and Pritchard D K (1980)
Comparative Risks of Electricity Production Systems, Health and Safety Executive Research Paper No 11, HMSO.

Collinson and Dowie (1980)
'Concepts and Classifications', Unit 1 of U201, *Risk: a second level University Course*, Open University Press, Milton Keynes.

Combs B and Slovic P (1979)
'Newspaper coverage of causes of death', *Journalism Quarterly*, 56(4), p 837–843.

Committee on Safety of Medicines (1985a)
'CSM Update' *British Medical Journal*, vol 291, p 1269.

Committee on Safety of Medicines (1985b)
'Adverse Reactions Working Party Report', Unpublished.

Culyer A J (ed) (1983)
Health Indicators, Martin Robertson.

Denning L J (1979)
The Discipline of Law, Butterworths.

de Dombal, F T *et al* (1974)
'Human and Computer-aided Diagnosis of Abdominal Pains: Further Report with Emphasis on Performance of Clinicians', *British Medical Journal*, vol 1, p 376/380.

Dowie J and Lefrere P (eds) (1980)
Risk and Chance, Open University Press.

Drug Surveillance Research Unit (1983)
PEM News, No 1.

English T A H *et al* (1984)
'The UK Cardiac Surgical Register, 1977–1982', *British Medical Journal*, vol 289, p 1205–1208.

Eraker S A and Sox H C (1981)
'Assessment of Patient's Preferences for Therapeutic Outcomes', *Medical Decision Making*, vol 1, p 29–39.

Evans R W *et al* (1984)
The National Heart Transplantation Study: Final Report, Seattle, Battelle Human Affairs Research Centre.

Fischoff B *et al* (1981)
Acceptable Risk, Cambridge University Press.

Gafni A and Torrance G W (1984)
'Risk Attitudes and Time Preference in Health', *Management Science*, vol 30, No 4, p 440–451.

Girdwood R H (1974)
'Death after taking Medicaments', *British Medical Journal*, vol 1, p 501–504.

Green C H and Brown R A (1978)
'Counting Lives', *Journal of Occupational Accidents*, vol 2, p 55–70.

Griffin J P and Weber J C P (1985)
'Voluntary Systems of adverse reaction reporting – Part I', *Adv Drug React Ac Pois Rev*, vol 4, p 213–230.

Griffin J P and Weber J C P (1986)
'Voluntary Systems of adverse reaction reporting – Part II, *Adv Drug React Ac Pois Rev*, vol 5, No 1.

Grist D R (1977)
Individual Risk – A compilation of Recent British Data, United Kingdom Atomic Energy Authority Report SRD R125.

Guarino R A (1979)
'Patient Package Inserts', *Food, Drug and Cosmetic Law Journal*, 34, p 113–120.

Henderson L (1977)
New England Journal of Medicine, vol 296, p 448.

Holroyd F and Collings S (1980)
'The Calculated Risk', Unit 2 of U201, *Risk: a second level University Course*, Open University Press, Milton Keynes.

Inman W H W (1984)
'Risks in Medical Intervention: Balancing Therapeutic Risks and Benefits', in *PEM News*, No 2, Drug Surveillance Research Unit, University of Southampton.

Inman W H W (1985)
'Side Effects of Drugs Essay 1985', in *PEM News*, No 3, Drug Surveillance Research Unit, University of Southampton.

International Commission on Radiological Protection (1977)
'Problems involved in developing an index of harm', Ann ICRP, 1, (4), (ICRP Publications No 27) Oxford: Pergamon Press.

Joint Formulary Committee (1985)
'Adverse Reactions to Drugs', *British National Formulary*, London.

Jombert P and Lasagna L (1975)
'Patient Package Inserts: Nature, Notions and Needs', *Clinical Pharmacology and Therapeutics*, vol 18, p 507–513.

Jones D R and Akehurst R L (1980)
'Risk Assessment: An Outline', *The Institute of Mathematics and its Applications*, vol 16, p 252–258.

Kahneman D and Tversky A (1979)
'Prospect Theory: an analysis of Decision under Risk', *Econometrica*, vol 47, No 2, p 263–291.

Kahneman D and Tversky A (1982)
'The Psychology of Preferences', *Scientific American Journal*, 1, p 160–173.

Keown C, Slovic P and Lichtenstein S (1981)
'Attitudes of Lay-people, Physicians and Pharmacists Towards Seriousness and Need for Disclosure of Prescription Drug Side Effects', *Health Psychology*.

Keown C, Slovic P and Lichtenstein S (1984)
'Influence of information about side effects on perceived risk of prescription drugs', *Health Marketing Quarterly*, vol 1, p 111-123.

Kletz T A (1977)
'What risks should we run?', *New Scientist*, 74, p 320–322.

Knight F H (1933)
Risk, Uncertainty and Profit, Houghton, Mifflin, Boston.

Knox E G (1975)
'Negligible Risks to Health', *Community Health*, vol 6, p 244–251.

Laurence D R and Black J W (1978)
The Medicine you Take, Fontana.

Lewis T L T (1980)
British Medical Journal, i, vol 295.

Lichtenstein S and Newman J R (1967)
'Empirical Scaling of Common Verbal Phrases Associated with

Numerical Probabilities', *Psychronomic Science*, 9, p 563–564.

Lichtenstein S, Slovic P, Fischoff B, Layman M and Combs B (1978)
'Judged Frequency of Lethal Events', *Journal of Experimental Psychology*, vol 4, p 551–578.

Mazis M, Morris L and Gordon E (1978)
'Patient attitudes about two forms of printed Oral Contraceptive Information', *Medical Care*, vol 16, p 1045–1054.

McNeil B J, Weichselbaum R and Pauker S G (1978)
'Fallacy of the five year survival in lung cancer' *New England Journal of Medicine*, vol 299, p 1397–1401.

McNeil B J, Weichselbaum R and Pauker S G (1981)
'Tradeoffs between Quality and Quantity of Life in Laryngeal Cancer', *New England Journal of Medicine*, vol 305, p 982–987.

McNeil B J, Pauker S G, Sox H C and Tversky A (1982)
'On the Elicitation of Preferences for Alternative Therapies', *New England Journal of Medicine*, 306, p 1259–1262.

Murphy A H and Winkler R L (1974)
'Probability forecasts: A survey of National Weather Service Forecasters', *Bulletin of the American Metereological Society*, 55, p 1449–1453.

National Consumer Council (1983)
Patients Rights, HMSO.

Office of Health Economics (1976)
Anaesthesia, OHE, London.

Office of Health Economics (1984)
Compendium of Health Statistics, 5th Edition, OHE, London.

Office of Population Censuses and Surveys (1985)
Hospital Inpatient Enquiry 1983.

Parker Hon Mr Justice (1978)
The Windscale Enquiry, Report presented to the Secretary of State for the Environment. HMSO.

Pochin E E (1975)
'The Acceptance of Risk', *The British Medical Bulletin*, 31, p 184–190.

Pochin E E (1981)
'Quantification of risk in medical procedures', *Proc Royal Society of London*, A376, p 87–107.

Pochin E E (1982)
'Risk and Medical Ethics', *Journal of Medical Ethics*, vol 8, p 180–184.

Prescott-Clarke P and Mostyn B J (1980)
'Public Attitudes towards the Acceptability of Risks', in, Prescott-Clarke P (1982)
'Public Attitudes towards Industrial, Work-related and Other Risks', Social and Community Planning Research, London.

Reissland J and Harris V (1979)
'A scale for measuring risks', *New Scientist* (13 September), p 80–81.

Lord Rothschild (1978)
'Risk', *The Listener*, 100, p 13–16.

Rosser R M and Watts V C (1972)
'The measurement of Hospital Output', *International Journal of Epidemiology*, 1, p 361–367.

The Royal Society (1983)
Risk Assessment: Report of a Royal Society Working Party, Royal Society, London.

Shakespeare W (1594)
The Merchant of Venice.

Sidaway *v* Gov of Bethlem Royal Hospital (HL(E))
The Weekly Law Reports, 8 March 1986, p 480–502.

Simpson, Chamberlain and Gravelle (1978)
'Choice of Screening Tests', *Journal of Epidemiology and Community Health*, 32, p 166–170.

Speirs C J *et al* (1984)
'Demography of the UK adverse reactions register of spontaneous reports', *Health Trends*, vol 16, p 49–52.

Speirs C J (1986)
'Prescription-related adverse reaction profiles and their use in risk-benefit analysis' in D'Arcy and Griffin (eds), *Iatrogenic Diseases*, (Third Edition), Oxford University Press.

Starr C (1969)
'Social Benefit Versus Technological Risk', *Science*, vol 165, p 1232–1238.

Teeling Smith G (1982)
Adverse Reactions and the Community, Office of Health Economics.

Teeling Smith G (1985)
Measurement of Health, Office of Health Economics.

Torrance G W (1986)
'Measurement of health state utilities for economic appraisal', *Journal of Health Economics*, vol 5, p 1–30.

Tversky A and Kahneman D (1974)
'Judgement Under Uncertainty: Heuristics and Biases', *Science*, vol 155, p 1124–1131.

Tversky A and Kahneman D (1981)
'The Framing of Decisions and the Psychology of Choice', *Science*, vol 219, p 453–458.

Urquhart J and Heilman K (1984)
Risk Watch, Kindler Verlag.

US Federal Register (1979)
'Prescription Drug Products: Patient Labelling Requirements' DHEW, FDA, 21CFR, PART 203, No 79N–0186 (July 6), 40016–40041.

von Neumann J and Morgenstern O (1947)
Theory of Games and Economic Behaviour, Princeton University Press.

Weinstein M C and Stason W B (1977)
'Foundations of Cost-Effectiveness Analysis for Health and Medical Practices', *New England Journal of Medicine*, vol 296, p 716–721.

Weinstein M C *et al* (1980)
Clinical Decision Analysis, W B Saunders Co, Philadelphia.

Wells N (1984)
Childhood Vaccination, Office of Health Economics.

Williams A (1985)
'Economics of Coronary Artery Bypass Grafting', *British Medical Journal*, vol 291, p 326–329.

Wilson R (1979)
'Analysing the daily risks of life', *Technology Review*, 81 (4), p 40–46.

OHE PUBLICATIONS

Studies of Current Health Problems ISSN 0473 8837

64 Scarce Resources in Health Care 60p
65 Schizophrenia 60p
66 Dementia in Old Age 60p
67 Huntington's Chorea 60p
68 Leukaemia: towards control 60p
69 Suicide and Deliberate Self-harm 60p
70 Alcohol: reducing the harm 60p
71 Hip Replacement £1.00
72 Medicines, Health and the Poor World £1.50
73 Coronary Heart Disease £1.00
74 Pharmaceutical Innovation £1.00
75 Understanding the NHS in the 1980s £1.50
76 Childhood vaccination – current controversies £1.00
77 Measurement of Health £1.00
78 Back Pain £1.00
79 Cystic Fibrosis £1.00
80 Crisis in Research £1.00
81 Health Expenditure in the UK £1.00
82 Health The Politicians Dilemma £1.00
83 Mental Handicap £1.50

OHE Briefings

Renal Dialysis 30p
The Effects of Prescription Charges 30p
Trends in European Health Spending 30p
Disability in Britain – the process of transition 30p
Sickness Absence: a review 30p
Accidents in Childhood 30p
Doctors, Nurses and Midwives in the NHS 50p
Medicine and the Quality of Life 50p
Ill in Europe 50p
Keep on Taking the Tablets? 50p
The Politics of Prescribing 50p

Reports on OHE Symposia

Health Research in England: a topic for debate £2.00
The Second Pharmacological Revolution £7.50
Measuring the Social Benefits of Medicine £7.50
A New NHS Act for 1996 £1.50
Pharmaceuticals among the Sunrise Industries £19.95
Health, Education and General Practice £1.50

Pharmaceutical Monographs

A Question of Balance: benefits and risks of medicines £1.50
The Consumer Movement, Health and the Pharmaceutical Industry £2.00
Issues in Development: a guide £2.00
The Future for Pharmaceuticals £2.00
Pharmaceuticals in Seven Nations £2.50

Compendia

Compendium of Health Statistics, 5th Edition: 1984 £15.00
Private Health Care in Britain £25.00

Research Monographs

Costs and Benefits of Regulating New Product Development in the UK Pharmaceutical Industry £5.00
The Needs of the Pharmaceutical Manufacturers from their Medical Departments in the 1990s £1.50
Patterns of European Diagnoses and Prescribing £5.00
Pharmaceuticals in Developing Countries: 1981–82 £2.00
Techniques of Health Status Measurement using a Health Index £2.50

Medicines: 50 years of progress *free*

Printed by White Crescent Press Ltd, Luton