Decision analysis in medicine

J G Thornton, R J Lilford, N Johnson

Introduction

For many straightforward medical problems any well trained doctor will make a good decision. Sometimes the correct course of action is unclear, however, and without help doctors and patients may make poor decisions because of a failure to consider probabilities correctly or to recognise the range of patients' values and weigh these correctly. Wrong decisions are made as a result of well recognised biases,¹ and one way of avoiding these biases and clarifying the problem is decision analysis.^{2,3}

Decision analysis is a method for breaking complex problems down into manageable component parts, analysing these parts in detail, and then combining them in a logical way to indicate the best course of action. In North America decision analysis is taught in most undergraduate medical courses but it is rarely used in the United Kingdom and was omitted from a *BMJ* series on logic in medicine in 1987.⁴ With more emphasis than ever before being put on patient choice in the NHS the time is ripe for a change of heart on decision analysis and we hope to go some way to remedy this national neglect in this article.

Gambling on probabilities

Anyone who has played 21 (vingt et un, pontoon, blackjack) has wondered when they are dealt, say, a 10 and a seven whether they should stick (remain with their two cards) or twist (buy another card). At any particular stage of the game the probability of victory with a score of 17 is known to the expert. Drawing another card may improve the odds of success but runs the risk of total loss by pushing the score over 21. In some versions of the game twisting may increase the stake and the potential winnings.

A typical medical problem is similar. Consider a doctor and a patient who have a choice between continuing with a current treatment with a known cure rate or performing a risky procedure which may improve the cure rate but which carries a risk of immediate death. The correct course of action depends on the probabilities (the cure rates and risks of the operation) and the values placed on the possible outcomes. In medicine tests may be available to revise risks, but usually these will also have some cost, if only to delay the start of definitive treatment. Decisions are particularly complex when there are multiple outcomes, all with different values to the patient.

Gamblers perform better if they calculate the odds and combine them correctly with the possible winnings. That is why professional poker players beat amateurs in the long run. Doctors can also improve their performance by calculating the risks and incorporating values correctly.

Decision analysis is explicit, quantitative, and prescriptive. It forces decision makers to spell out the way decisions have been broken into their component parts and then recombined. Decision makers are compelled to measure, and put numerical values on, both key

uncertainties and the values of possible outcomes. Decision analysis aims at telling doctors what to do, not just describing what they do. There are four basic steps in a decision analysis.

- (1) Identify the decision problem.
- (2) Structure the decision problem over time.
- (3) Measure the uncertainties (probabilities and utilities) needed to fill in the structure.
- (4) Combine the uncertainties to choose a preferred course of action.

The following example, based on a full decision analysis to be published elsewhere,⁵ introduces the technique of classical decision analysis.

The problem

A 29 year old nurse, engaged to be married and planning a family, developed occult cervical cancer. The diagnosis was confirmed by a cone biopsy, which showed a moderately differentiated squamous cancer, invading 2 mm below the basement membrane and with lymphatic spread. The primary tumour was completely excised and the pathologist reported a wide margin of normal tissue around the tumour. What treatment should she choose?

We began by defining the problem. The treatment options lay between no further treatment and extended hysterectomy with lymphadenectomy. We ruled out simple hysterectomy since this could not remove tumour metastases, was unlikely to improve survival, and would automatically render her infertile. The possible outcomes considered were survival with fertility retained, immediate death from surgical complications, delayed death from cancer, and survival but with infertility.

The next stage was to structure the problem using a decision tree (fig 1). This is a flow diagram in which decisions and outcomes are represented in order with early events to the left and later events to the right. Decision points are represented by square nodes and points where outcomes occur by chance by round nodes. In figure 1 the left hand decision node repre-

FIG 1—Basic decision tree for microinvasive cancer of the cervix without probabilities and utilities

Institute of Epidemiology
and Health Services
Research, University of
Leeds, Leeds LS2 9JT
J G Thornton, senior lecturer
R J Lilford, professor
N Johnson, senior lecturer

Correspondence to:
Mr J G Thornton,
Department of Obstetrics
and Gynaecology, Clarendon
Wing, Leeds General
Infirmary, Leeds LS2 9NS.

BMJ 1992;304:1099-103

sents the choice between no further surgery and radical hysterectomy. The upper circle is a chance node representing the chance that the patient may die of cancer or survive in full health having retained her fertility if no surgery is performed. The other nodes are self explanatory. The order of events in this decision tree needs to be chosen with care. The chance outcome of "spread" or "no spread" obviously occurs biologically before operation but will become known only after the operation, when the excised specimen has been examined. It is therefore placed after the operation.

Our third step was to fill in the probabilities and utilities of each outcome. There was no meta-analysis of randomised trials, or even a single randomised trial, to provide these, so we had to base our estimates on observational studies just as we would have had to do if we were not using decision analysis. After a literature review we estimated surgical mortality as 5/1000, the likelihood of disease spread beyond the area of the cone biopsy as 2%, and the chance of cure by surgery if it had done so as 50%.³ In figure 2 the probability of each outcome (estimated from the literature) is represented in brackets beside each outcome as a number between 0 (will never occur) and 1 (certain to occur). All possible outcomes are included so that the probabilities at each chance node always add up to 1.

FIG 2—Decision tree for microinvasive cancer of the cervix with probabilities added

FIG 4—Hypothetical utility curves for life expectancy and money. The horizontal scale represents either life expectancy or money. If it represents life expectancy the dotted line at 45° represents the utilities of a person whose scale of values is linear: 25 years has exactly half the utility of 50 years. This is an unusual attitude. More often people are risk averse for life expectancy and the upper curve better reflects their values: here 15 years have half the utility of 50 years. Such a person would be indifferent between a certain 15 years' life and a 50/50 lottery between immediate death and 50 years of life. If the horizontal axis is taken to indicate money the lower curve represents the utilities of a risk seeking gambler

preferable to delayed death, but she also needed to know exactly where to place the intermediate states on her utility scale.

She first calculated the utility of infertile life by choosing between that and various gambles between fertile life and immediate death until she reached a level of indifference. It worked as follows. She was asked to imagine two doors, through one of which she had to go. Behind the left hand door there was no risk of death but she would be rendered infertile. Behind the right hand door she would encounter a 50% chance of fertile survival but also a 50% risk of death (fig 3). She chose the left hand door. The risks of death through the right hand door were decreased until a point was reached where she could not decide which door to select. This occurred when the risk of death through the right hand door was 5% and of fertile survival 95% (fig 3). This was the level of indifference. Our patient therefore valued survival with infertility as 0.95 on a scale where full health was valued 1 and immediate death valued 0. She performed a second similar lottery between delayed death from cancer and various chances of full health or immediate death and derived a utility for delayed death of 0.05.

There are alternative methods of measuring values, such as asking patients to mark health states on a linear scale, but, unlike the reference gamble, this method is not axiomatically correct. People avoid the extremes of the scale, and because they may not perceive the trade off inherent in the technique the values obtained in this way may be distorted. A better alternative makes use of natural underlying scales such as money or years of life. Unfortunately people's utilities for money and years of life are rarely linear. People are usually risk seeking or risk averse. For example, the gambler in our earlier example is likely to be risk seeking. A utility curve for such a gambler is shown in figure 4, where £50 has twice the utility of £35. People taking out insurance policies are by definition risk averse and will have utility curves of the shape of the upper left hand curve in fig 4.

Years of life expectancy is another frequently used underlying scale, but people tend to value the years immediately ahead more highly than those far in the

FIG 3—(Top) First lottery to measure the utility of infertility. (Bottom) Point at which the patient was unable to decide whether to risk death to avoid infertility (the level of indifference). This defines the utility of infertility as 0.95 on a scale where immediate death is valued 0 and fertile life as 1

This information enabled us to calculate the treatment with the lowest expected mortality. For our patient this was 0.02 with no further treatment and 0.015 ($0.005 + (0.02 \times 0.5)$) with radical hysterectomy. This does not, however, mean that radical hysterectomy is the best treatment because we have not considered all relevant factors—for example, the patient's preference for delayed death from cancer versus immediate death from surgery, the desire to avoid the morbidity of surgery, and above all the desire to conserve fertility. We need to measure the utilities of the outcomes in a way that will allow us to see what chance of one favoured outcome our patient will relinquish to obtain another favoured outcome. When we have such a measure we can combine these utilities with the probabilities in a logical fashion to calculate the treatment with the highest expected utility.

Measurement of utilities

The best method for measuring people's utilities is the basic reference lottery where the relative utilities of three health states are worked out together.^{6,7} Our patient needed to define the utility of four health states so two lotteries were needed. She ranked the health states as follows: the best was fertile life (with a utility of 1), the worst immediate death (with a utility of 0), with infertile life and delayed death rated intermediate. She did not find it difficult to rank infertile life as

future. This is another example of risk aversion, and utility scales must reflect this. The upper left hand curve in fig 4 represents such a curve when the horizontal scale is converted to years of life expectancy.

The non-linearity of monetary and life expectancy scales makes it impossible to use them without adjustment to calculate expected utilities in decision analysis. Either basic reference lotteries must be performed to measure the relevant utilities directly or utility curves for the relevant patient or population of patients must be derived from monetary or life expectancy lotteries.

Adding utilities to the decision analysis

Having measured the utilities we need to combine them with the probabilities to select a preferred course of action—that is, that with the greatest expected utility. We start by estimating the utility of each chance node, which is calculated as the weighted average of the utilities of its possible outcomes, where the weights are the probabilities of each outcome. The utility of the upper chance node in figure 5 is thus $(0.02 \times 0.05) + (0.98 \times 1.0) = 0.981$. Where there is a sequence of chance nodes in the tree we use the weighted utility of the distal chance node in calculating the expected utility of the proximal node. The utility of a decision node is the maximum of the utilities of its component branches since a rational decision maker should choose this strategy. It is clear that the expected utility of no further surgery (0.981) is greater than that of radical hysterectomy (0.936), and this is the option our patient should choose.

FIG 5—Decision tree for microinvasive cancer of the cervix with outcome utilities added and the expected utilities for each branch (dotted lozenges) and for each chance node (solid lozenges) calculated

The difference in expected utility between the different courses of action may not appear very great, but on this scale the difference 0.045 represents 4.5% of the value of the patient's entire life in full health. Moreover, if the axioms of expected utility theory are accepted by our patient (most people do agree that this is how they wish to make decisions), and if the probability and value estimates are the best possible, then it would be perverse to choose the course of lower utility, however small the difference.

Sensitivity analyses

The final part of a full decision analysis should include a sensitivity analysis, because conclusions depend on the probabilities and utilities used, and in real life we are rarely, if ever, certain what these are. In a sensitivity analysis each of the key probabilities and values is varied in turn within the range of reasonable uncertainty to test the robustness of the conclusion. Figure 6 shows a one way sensitivity analysis to show the effect of varying the utility of infertility. Each straight line on the graph represents the expected

utility of the relevant strategy at a range of levels of infertility utility. The strategy lines intersect at an infertility utility of 0.995; therefore above this value radical surgery is the preferred option while below it no further surgery is preferred. The point at which strategy lines intersect is called a decision threshold. This threshold will itself vary if other variables such as operative mortality and recurrence risk are changed.

FIG 6—A one way sensitivity analysis of the microinvasive cancer of the cervix decision to show the effect of varying the utility of infertility

The effect of changing more than one variable can be shown in a threshold analysis (fig 7). Here the decision threshold is plotted against the risk of recurrence and utility of infertility for three different operative death rates. For each patient the utility of infertility and probability of disease spread is plotted. If this point falls below and to the left of the relevant threshold line she should not undergo surgery and if above and to the right she should. The effect of varying utilities and probabilities can be seen at a glance. Decision trees and sensitivity and threshold analyses can easily become complicated and so computerised aids are widely used by serious practitioners.⁸

FIG 7—A threshold analysis of the microinvasive cancer of the cervix decision to show the effect of varying the utility of infertility, the probability of spread, and operative mortality simultaneously

In this example our conclusions were very sensitive to the value placed on retaining fertility. Doctors should therefore take great care to explore this particular issue with a patient. The example was inspired by a debate about the histological diagnosis of so called "microinvasive cancer." The analysis shows how treatment can be tailored to individual patients according to a range of histological and other criteria and that any attempt to treat patients according to fixed cut off criteria for histological diagnosis is doomed to failure. Decision analysis is seldom used in its full rigour for the treatment of individual patients. More often it helps to structure debate, and it has been used to shed light on some important medical controversies. Some examples are listed in the box. Decision analysis has also been used for some novel purposes, such as the design of randomised controlled trials,^{22 23} and the lottery method has been used to clarify ethical

Published decision analyses

- Amniocentesis for prenatal diagnosis⁹
- Management of the term breech¹⁰
- Chemotherapy for node negative breast cancer¹¹
- Anticoagulation for deep vein thrombosis in pregnancy¹²
- Oestrogen replacement in the menopause¹³
- Elective caesarean section¹⁴
- Coronary artery bypass surgery¹⁵
- Antibiotic treatment for sore throats¹⁶
- Management of appendicitis¹⁷
- Occult blood testing for bowel cancer¹⁸
- Management of intracranial aneurysms¹⁹
- Timing of prenatal diagnosis²⁰
- Management of ovarian cancer²¹

problems.²⁴ It is closely allied to cost-utility studies used for resource allocation decisions.

Why prefer decision analysis to global decision making?

The extreme alternative to formal decision analysis is for doctors and patients to decide by intuition and to make a single global analysis of the whole problem. People deciding this way often make suboptimal decisions. These may result from inadequate data, and decision analysis cannot help here except to indicate the need for careful review of the evidence.

Ideally each probability estimate used in making a decision should be derived from published data; in the case of occult cervical tumours we found many relevant articles. Ideally the results from large randomised controlled studies should be used and they should be combined using formal techniques such as meta-analysis.²² Often such studies are unavailable, and even when they are judgment is required to extrapolate results from one time and place to another. Semi-objective probability estimates often have to be used in which probabilities obtained from published data are modified according to local circumstances or changes in practice. For example, we adjusted the published estimate of operative mortality for occult cervical cancer downwards to take account of improvements in surgical technique, anaesthetics, and intensive care since the studies were undertaken. One of the most common and least justified criticisms of decision analysis is that the need to make such revisions of probability estimates invalidates the technique. These adjustments invalidate decision analysis no more than they do conventional intuitive decision making. The latter is also based on probabilities, which are no more accurate for not being made explicit. Indeed, the process of making probabilities explicit is a reason to use rather than abandon decision analysis, since this exercise exposes the source of disagreement about treatment policy. Because it is transparent, decision analysis encourages a comprehensive review of published data.

Probabilities also need to be adjusted for individual patients in the light of specific test results; the probability of an event is often based not on one but on many items of information. The usual technique for working out the probability of an event, by combining prior probability (prevalence) and the result of all these tests, is Bayes's theorem. The archetypal example of the use of Bayes's theorem in clinical medicine is in the diagnosis of abdominal pain described by de Dombal in 1972.²⁵ There is evidence that this method improves clinical care by reducing delay in the surgical treatment of gangrenous organs while at the same time reducing the number of negative laparotomies.²⁶ The probabilities of different diagnoses produced in this way are eminently suitable for including in decision trees and hence for decision analysis.

In the absence of data from meta-analysis of good trials and correct Bayesian revision of risks humans use a number of heuristics (systems of reasoning) to estimate subjective probabilities. These often lead to predictable and well recognised biases.¹ For example, we often estimate the frequency of an event from its ease of recall. Vivid and recent events are overestimated as a result. A surgeon who has just had a death from treatment complications and a patient whose friend died postoperatively will both overestimate treatment risks. Another heuristic is to use the degree of similarity of a pattern of observations to an event to estimate the likelihood of that event. For example, if an unknown fruit looks like an apple it probably is an apple, but if the fruit comes from a country where apples do not grow—that is, the prior odds of it being

an apple are low—the “representativeness” heuristic will lead to an overestimate. Finally the “anchoring” heuristic describes the way in which people typically make a prior estimate of an event occurring. When further information becomes available they adjust this prior risk upwards or downwards. Biases occur if the prior estimate was wrong and because adjustments in the light of new information are typically insufficient and often not made at all. For example, doctors may simply ignore test results that do not fit in with their preconceived ideas.

Finally, decision analysis provides a method to put large amounts of information together. Intuitive decision makers may claim that they are incorporating more complex patient preferences of intermediate outcomes in the decision, but analysis of how they actually make decisions reveals that they also make major simplifying assumptions.²⁷ Many of these simplifications are made necessary simply by the limitations of human short term mental capacity. It is impossible to keep more than a few facets of a decision in the forefront of the mind at any one moment. Decision analysis aids the overworked brain by separating the components of a decision so that they can be analysed separately.

Nevertheless, doctors often still resist decision analysis. They may feel that it is unnecessary for the many well defined everyday decisions where either the consequences of a wrong diagnosis are unimportant or the correct course of action is clear to everyone. This is no argument for not using it for the less common difficult decisions. It is not popular with those doctors who deal with poorly defined conditions—irritable bowel syndrome or premenstrual tension, for example. There are no objective methods to confirm or refute such “I say so” diagnoses, and the outcomes are generally good whatever is done. For this reason decision analysis is perceived as having little to add. Since it emphasises the way that values and probabilities underlie decisions it is threatening to those who like to work with certainties. Finally, many clinicians feel that utility functions are unnecessarily precise. Is infertility really 0.95 of full health? These problems have not all been resolved and undoubtedly explain why decision analysis has not been more widely used.²⁸

Conclusion

The language and methods of decision analysis and more specifically of expected utility theory can change how we think. The discovery that there is a specific mathematical function (expected utility) which measures the benefits of a course of action is a revelation to many when they first come across it. We hope that this article inspires some practitioners. Decision analysis has been widely used in business for years and has entered the mainstream of medical thinking in North America and more recently Australia and New Zealand. It is incorporated in the medical curriculum in centres as far apart as Hamilton, Ontario, and Dunedin in the South Island of New Zealand. We believe that doctors in Europe may love it or hate it but cannot ignore it.

1 Tversky A, Kahnemann D. Judgement under uncertainty: heuristics and biases. *Science* 1974;185:1124-31.

2 Weinstein MC, Fineberg HV, eds. *Clinical decision analysis*. Philadelphia: Saunders, 1980.

3 Pauker SG, Kassirer JP. Decision analysis. *N Engl J Med* 1987;316:250-8.

4 Philips C, ed. *Logic in medicine*. London: BMJ, 1988.

5 Johnson N, Lilford RJ, Jones SE, McKenzie L, Billingzly P, Songane F. Using decision analysis to calculate the optimum treatment for micro-invasive cervical cancer. *Br J Cancer* 1992;65:717-22.

6 Keeny RL, Raiffa H. *Decisions with multiple objectives: preferences and value tradeoffs*. New York: Wiley, 1976.

7 Lindley DV. The role of utility in decision making. *J R Coll Phys Lond* 1975;9:225-30.

- 8 McNamee P, Celona J. *Decision analysis for the professional with Supertree*. Redwood City, California: Scientific Press, 1987.
- 9 Pauker SP, Pauker SG. Prenatal diagnosis: a directive approach to genetic counselling using decision analysis. *Yale J Biol Med* 1977;50:275-89.
- 10 Bingham P, Lilford RJ. Management of the selected term breech presentation: assessment of the risks of selected vaginal delivery versus Cesarean section for all cases. *Obstet Gynecol* 1987;69:965-78.
- 11 Hillner BE, Smith TJ. Efficacy and cost effectiveness of adjuvant chemotherapy in node negative breast cancer. *N Engl J Med* 1991;324:160-8.
- 12 Klein K, Pauker SG. Recurrent deep venous thrombosis in pregnancy: analysis of the risks and benefits of anticoagulation. *Med Decision Making* 1981;1:181-202.
- 13 Elstein AS, Holzman GB, Ravitch MM, Metheny WA, Holmes MM, Hoppe RB, et al. Comparison of physicians decisions regarding estrogen replacement for menopausal women and decisions derived from a decision analytic model. *Am J Med* 1986;80:246-58.
- 14 Feldman GB, Freiman JA. Prophylactic caesarean section at term? *N Engl J Med* 1985;312:1264-7.
- 15 Weinstein MC, Pliskin JS, Stason WB. Coronary artery bypass surgery: decision and policy analysis. In: Bunker JP, Barnes BA, Mosteller F, eds. *Costs, benefits and risks of surgery*. Oxford: Oxford University Press, 1977.
- 16 Tompkins RK, Burnes DC, Cable WE. An analysis of the cost effectiveness of pharyngitis management and acute rheumatic fever prevention. *Ann Intern Med* 1977;86:481-92.
- 17 Neutra R. Indications for the surgical treatment of suspected acute appendicitis. A cost effectiveness approach. In: Bunker JP, Barnes BA, Mosteller F, eds. *Costs, benefits and risks of surgery*. Oxford: Oxford University Press, 1977.
- 18 Neuhauser D, Lewicki AM. What do we gain from the sixth stool guaic? *N Engl J Med* 1975;293:226-8.
- 19 Van Crevel H, Habbema JDF, Braakman R. Decision analysis of the management of incidental intracranial saccular aneurysms. *Neurology* 1986;36:1335-9.
- 20 Lilford RJ. Trade off between gestational age and miscarriage risk of prenatal testing: does it vary according to genetic risk? *Lancet* 1990;336:1303-5.
- 21 Simes RJ. Treatment selection for cancer patients: application of statistical decision theory to the treatment of advanced ovarian cancer. *J Chron Dis* 1985;38:171-86.
- 22 Sacks HS, Berrier J, Reitman D, Ancona-Berk VA, Chalmers TC. Meta-analysis of randomised controlled trials. *N Engl J Med* 1987;316:450-5.
- 23 Lilford RJ. Clinical trial numbers. *Lancet* 1990;335:483-4.
- 24 Thornton JG, Bryce FC, Bhabra K. Withholding intensive care from premature babies. *Lancet* 1988;ii:332-3.
- 25 De Dombal FT. Computer aided diagnosis of acute abdominal pain. *BMJ* 1972;ii:9-13.
- 26 Adams ID, Chan M, Clifford PC, Cooke WM, Dullos V, de Dombal FT, et al. Computer aided diagnosis of abdominal pain: a multicentre study. *BMJ* 1986;293:800-4.
- 27 Elstein AS, Bordage G. Psychology of clinical reasoning. In: Dowie J, Elstein A, eds. *Professional judgement. A reader in clinical decision making*. Cambridge: Cambridge University Press 1988:109-29.
- 28 Balla JJ, Elstein AS, Christensen C. Obstacles to acceptance of clinical decision analysis. *BMJ* 1989;298:579-82.

(Accepted 29 December 1991)

Basic Molecular and Cell Biology

Gene regulation

David S Latchman

This is the first of two articles updating the series "Basic Molecular and Cell Biology" published in 1987. The articles will be published in a new edition of the book of the series later this year.

That the expression of human genes must be a highly regulated process should be clear to anyone who has ever dissected a human body. The vast range of different tissues and organs differ dramatically from each other and they all synthesise different proteins—haemoglobin in red blood cells, myosin in muscle, albumin in the liver, and so on. Moreover, with few exceptions all these different cell types contain the same sequence of DNA, which encodes all these different cell proteins, and this DNA is also identical to the DNA in the single celled zygote, from which all these different cells arise during embryonic development. Clearly, therefore, some process of gene regulation must operate to decide which genes within the DNA will be active in producing proteins in each cell type.

Levels of gene regulation

A number of stages exist between the DNA itself and the production of a particular protein (fig 1).¹ Thus the DNA must first be transcribed into a primary RNA transcript, which is subsequently modified at both ends by the addition of a 5' cap and a 3' tail of adenosine residues. Moreover, within this primary transcript, the RNA sequences which actually encode the protein are not present as one continuous block. Rather they are broken up into segments (exons) which are separated by intervening sequences (introns) that do not contain any protein coding information. As these introns interrupt the protein coding region and would prevent the production of an intact protein they must be removed by the process of RNA splicing² before the mature messenger RNA can be transported from the nucleus to the cytoplasm and translated into protein.

Clearly each of these stages is a potential point at which gene expression could be regulated, and there is evidence that several of them are actually used. Thus, for example, the production of many new proteins in the egg immediately after fertilisation and the start of embryonic development depends on the translation into protein of fully spliced, messenger RNAs that pre-existed in the cytoplasm of the unfertilised egg but

whose translation was blocked before fertilisation. This form of gene regulation is known as translational control. Similarly, by splicing the protein coding regions (exons) of a single primary transcript in different combinations two or more different mRNAs encoding different proteins in different tissues can be produced. This process of alternative splicing³ is well illustrated in the single gene that encodes both the calcium modulating hormone, calcitonin, and the

FIG 1—Stages in gene expression which could be regulated

Division of Molecular Pathology, University College and Middlesex School of Medicine, London W1P 6DB
David S Latchman, professor and head of division

BMJ 1992;304:1103-5