Decision analysis for medical managers

J G Thornton, R J Lilford

This is the twelfth in a series of articles dealing with issues arising as clinical practitioners increasingly take on managerial roles. The series is edited by Jenny Simpson, chief executive, British Association of Medical Managers. A book of the series is now available from the BMJ Bookshop.

Case study

The obstetricians at "Enterprise Hospital Trust" in the northern city of Yeeds are trying to persuade the health authority to buy serum screening for Down's syndrome for all pregnant women in the district.¹ The director of purchasing was about to approve this when she received a letter from Lady Pressure, chair of the community health council. Lady Pressure's grandchild had just died of congenital toxoplasmosis, and she wondered why this screening was not offered in Yeeds. She enclosed literature from the Toxoplasmosis Trust calling for a screening programme, and reminded the manager that screening had been offered in France for years. Why was Yeeds so backward? Finally, the manager herself was aware that the recent identification of the cystic fibrosis gene made screening for that disease technically simple.

The right decision in this case is not obvious. It is tempting to choose the cheapest programme, but perhaps a slightly more expensive one will produce greater health gain. This itself is difficult to measure, and different programmes do not simply save different numbers of lives. How can we compare prevention of the birth of a baby with Down's syndrome with the birth of a baby with cystic fibrosis? Each programme will also have other "costs": they will cause some miscarriages and make quite a few parents anxious.

Typically, managers resolve such conflicts by reference to previous practice and to what others do and by subjectively evaluating the claims of the various pressure groups. The problem with such political methods is that "he who shouts loudest" is likely to win, without necessarily being the most deserving. Sometimes decision makers will want to stand back and attempt a more thoughtful analysis.

Decision analysis

Decision analysis is a way of doing this by analysing the benefits and harms systematically, so that the trade offs are explicit.² Business people often use it,³ and much of the literature refers to their problems.⁴ An increasing amount of published data concerns health management problems such as neonatal intensive care of very low birthweight infants⁵ or colposcopy.⁶ The conclusions are often surprising. The marginal cost of preventing one death from bowel cancer using the recommended policy of measuring a sixth faecal occult blood sample was estimated at US \$47 million in 1975.⁷ Screening for ovarian cancer increases women's life expectancy by an average of one day.⁸

We live in a political world, and wise managers will rarely decide solely on the basis of decision analysis. Nevertheless, it will inform their final decision and provide a powerful means to defend it. Those who dislike formal methods must recognise that there is little rational alternative within a planned health care system. Individual planners may wish to follow their own intuitions when they make decisions, but they can hardly expect others to accept that method. A market, with many consumers deciding how to spend their

health care money, would be sensitive to more factors than those which comprise decision analysis,⁹ but it may not provide the equity which current health care systems seek.

Decision analysis is the application to decision making of the reductionist approach that has been so successful elsewhere in science.

CLINICAL DECISIONS

Clinical decision analysis usually guides treatment for an individual. Consider a pregnant woman wondering whether to undergo amniocentesis for the prenatal diagnosis of spina bifida. After measurement of serum α fetoprotein and an ultrasound scan her residual risk has been estimated to be 1 in 200.¹¹ However, amniocentesis carries a small risk of miscarriage (fig 1) and the correct course also depends on the disutility to her of having an affected child or losing a normal pregnancy. We measure these by a series of lotteries (fig 2). For the individual to whom a 25% chance of spina bifida and pregnancy termination is equal, we say that the disutility of termination is 0.25 on a scale where a normal baby is 0 and one with spina bifida 1. The course of minimum expected disutility is calculated by multiplying probabilities and disutilities.

"Wise managers will rarely decide solely on the basis of decision analysis."

In our examples we measure disutility rather than utility as most people find this easier when the bad outcomes of miscarriage or handicap are less frequent than the good ones. We use disutility rather than the more euphonious term cost to avoid confusion with monetary costs.

ANALYSIS OF MANAGEMENT DECISIONS

The above example is concerned with helping an individual, but managers can use the same techniques

FIG 1—Decision tree for amniocentesis for spina bifida. The term "normal" baby is used to indicate an unaffected child. Although many other abnormalities may occur, these will be equally distributed between the two arms of the decision tree. The expected disutility of no amniocentesis is disutility of spina bifida $\times$ probability of spina bifida $= 0.005 \times 1 = 0.005$. The expected disutility of amniocentesis is (disutility of procedure related miscarriage $\times$ probability of miscarriage) $+ (disutility of termination \times probability of termination) = (0.25 \times 0.01) + (0.25 \times 0.005) = 0.00375$. The expected disutility of undergoing amniocentesis is thus less than that of not doing so, and the patient should choose amniocentesis.

Department of Obstetrics and Gynaecology, University of Leeds, Leeds LS2 9LN
J G Thornton, honorary consultant

Institute of Epidemiology and Health Services Research, University of Leeds, Leeds LS2 9LN
R J Lilford, chairman

Correspondence to: Dr Thornton.

BMJ 1995;310:791-4

FIG 2—The lotteries for the spina bifida decision. For simplicity we assume that the disutility of procedure related miscarriage is equal to that of termination. A second lottery could be performed if these were significantly different

to decide for groups. It is easy enough to calculate the relevant probabilities, but difficult to include utilities because there is no generally accepted way of measuring these for groups, and there are few empirical data. Nevertheless, managers have to make decisions, and in doing so an assessment of population utilities is implied—whether or not it is made explicit.

Let us perform population decision analyses for screening for Down's syndrome, congenital toxoplasmosis, and cystic fibrosis. To keep things simple, we focus on the outcomes affected by screening, and ignore background rates of disability and miscarriage. In all three examples we refer to a hypothetical population of 100 000 pregnancies. We consider first the probabilities of each outcome. Next we compare the disutilities of these outcomes and combine these

with the probabilities. Lastly, we discuss monetary costs.

Probabilities of various outcomes

SCREENING FOR DOWN'S SYNDROME

Without screening 100 babies with Down's syndrome would be born.¹² Assuming an 80% uptake for the screening test and a 75% uptake of amniocentesis with risks of over 1 in 250, 80 000 women would undergo screening and 4000 would screen positive; 3000 would undergo amniocentesis, of whom 30 would miscarry a normal baby. In the process 40 Down's syndrome babies would be detected and the pregnancies aborted (fig 3).

CONGENITAL TOXOPLASMOSES SCREENING

A screening programme would involve serial maternal blood tests in pregnancy to identify women who seroconvert, fetal blood sampling to confirm fetal infection, and either antibiotic treatment or abortion of infected fetuses. At best, screening would prevent the birth of 10 disabled babies, for the disutility of 10 terminations and five miscarriages.¹³ At worst, the birth of only one disabled baby would be prevented for the disutility of 40 terminations and 12 miscarriages.¹⁴

"Instead of just measuring anxiety, psychologists will in future ask what it means in benefits that should be forgone."

This uncertainty is one reason why toxoplasmosis screening is not generally offered in the United Kingdom. For the present we will make a "best guess" and assume that five miscarriages related to the procedure and 20 terminations would be caused, to prevent the birth of five babies with congenital toxoplasmosis (decision tree not shown).

CYSTIC FIBROSIS SCREENING

Without screening some 40 babies would be born with cystic fibrosis. In a typical programme women would be offered carrier testing first and their partners tested if the woman was positive. Carrier couples would have a one in four risk of having an affected child and would be offered invasive testing. Typically 80 women would undergo amniocentesis or chorionic villus sampling, one would miscarry, and 20 babies with cystic fibrosis would be identified (decision tree not shown).¹⁵

To decide whether such screening programmes give a net health gain we need to know not only the risks but also the relative disutilities to typical women of the loss of a normal baby and the birth of a child with Down's syndrome, congenital toxoplasmosis, or cystic fibrosis. We also need to measure the anxiety caused by the programme.

DOWN'S SYNDROME DISUTILITIES

Samples of women have performed the relevant lotteries,^{16 17} and miscarriage or termination typically has a disutility of about 0.3 on a scale where full health is 0 and Down's syndrome is 1 (figs 4 and 5). The expected disutility of a screening programme is thus:

70 miscarriages or terminations $\times 0.3 = 21$;
that of not screening is:

40 Down's $\times 1$ (disutility of Down's) = 40.
Screening provides a net gain of 17 "utility units."

One source of error in this calculation is the accuracy of the population utilities used. We have used the

FIG 3—Decision tree for Down's syndrome screening for a population of 100 000 pregnancies

FIG 4—Lottery for measuring the relative disutilities of termination or miscarriage and Down's syndrome

medians of possibly unrepresentative samples because these are all that are available. As population health utility surveys, such as the Oregon experiment,¹⁸ become more widespread, representative mean utilities may become available and would be preferable.

Decision analysis is unashamedly utilitarian. It cannot select the correct course of action for people who wish to follow other ethical principles. Those who follow the rule that all human life is sacred from conception onwards will not agree with us, and our analysis ignores their views except in so far as the socially derived utilities take them into account.

TOXOPLASMOSIS UTILITY

Like the risks, the disutility of having a child with congenital toxoplasmosis is less well defined than that of Down's syndrome. The disease varies in severity from mild to severe developmental delay, and the more severely affected cases may not survive long. Toxoplasmosis may therefore be of less disutility than Down's syndrome, which has a typical life expectancy of 55 years. For our illustration we will assume that toxoplasmosis has only twice the disutility of miscarriage or termination—that is, that people would be indifferent between abortion and a 50% risk of congenital toxoplasmosis (fig 6). This corresponds to a disutility of 0.6 (fig 5). The "best guess" toxoplasmosis screening programme thus has an expected disutility of:

25 miscarriages or termination $\times 0.3 = 7.5$;
not screening a disutility of
5 affected babies $\times 0.6$ (toxoplasmosis disutility) = 3.

Screening, far from increasing utility, decreases it by 4.5 units.

If the figures on which an analysis is based are uncertain it may be wise to repeat the analysis with different but still plausible risks to see if the best decision changes on sensitivity analysis. A toxoplasmosis screening programme might only cause 15 miscarriages or terminations for every 10 affected births prevented. In that case screening would have a disutility of $15 \times 0.3 = 5$ units, while not screening would have a disutility of $10 \times 0.6 = 6$ units. It seems that even under the most favourable assumptions toxoplasmosis screening adds only marginal value to the community.

CYSTIC FIBROSIS UTILITY

There are few data on the disutility of cystic fibrosis, although it is probably perceived as less severe than Down's syndrome and perhaps, because the brain is not affected, as less severe than toxoplasmosis. Let us assume that women on average would be indifferent between a 60% risk of cystic fibrosis and abortion or miscarriage (fig 7), corresponding to a cystic fibrosis disutility of 0.45 (fig 5). The expected disutility of not screening is thus:

20 cystic fibrosis $\times 0.45 = 9$;
that of screening is:

21 miscarriages or terminations $\times 0.3 = 7$.

The screening programme increases utility by 2 units.

The analysis so far suggests that screening for Down's syndrome and for cystic fibrosis increases expected utility, the former more so, while toxoplasmosis screening decreases utility.

THE ANXIETY CAUSED BY SCREENING

Much public concern about screening is based on the

FIG 5—Diagram of the disutilities used in the present analyses. Only the disutility of miscarriage or termination and Down's syndrome are based on empirical data. The disutilities of toxoplasmosis and cystic fibrosis and the anxiety of being screen positive seem plausible but have not been formally measured

FIG 6—Lottery for measuring the relative disutility of termination or miscarriage and toxoplasmosis

anxiety it causes to many, compared with the good it might do for a few. Even a verbal offer to screen will induce some anxiety, and being screened positive, even if ultimately diagnosed as negative, will cause still more.¹⁹ This anxiety should be debited to the screening programme. The first step is to measure the extent and duration of anxiety, and the second task is to put it on to our utility scale. In the absence of empirical data let us assume that the anxiety caused by being positive for Down's syndrome by screening is such that 1000 women made anxious may be traded off against the prevention of one child with Down's syndrome being born (fig 8), and that the anxiety of being a cystic fibrosis carrier is half this because of the lesser severity of the disease. The analysis would go as follows.

In 100 000 pregnancies 4000 women would be rendered anxious because of a "screen positive" test result for Down's syndrome. This is equivalent to $4000 \times 0.001 = 4$ Down's syndrome units. This is added to the other disutilities of screening, 23 units, to make a total disutility of 27 units. This is still outweighed by the prevention of 40 units by screening.

Screening for cystic fibrosis would render 5000 women anxious ($5000 \times 0.0005 = 2.5$ units) to be added to the 7 units from miscarriage or termination, making a total disutility of 9.5 units from screening to prevent 9 units from disability. The decision is finely balanced for cystic fibrosis screening, with the conclusion being sensitive to the anxiety of being a carrier, something we do not pretend to know. However, as more managers use decision analysis, these issues will be thrown into sharper focus, and we anticipate that instead of just measuring anxiety, psychologists will in future ask what it means in benefits that should be forgone.

FIG 7—Lottery for measuring the relative disutility of termination or miscarriage and cystic fibrosis

FIG 8—Lottery for measuring the disutility of "screen positive" anxiety in terms of a Down's syndrome birth

Key point summary

- When doctors treat patients good outcomes must be traded off against bad, taking into account the values of the individuals involved. In management the perceptions of the community as a whole determine the values
- Decision analysis is a way to separate the treatments that add net value from those that do not, and to compare the relative value of beneficial treatments, on the assumption that it might not be possible to implement them all
- Decision analysis is a powerful means of defending decisions taken, and can be applied to individuals and groups

FINANCIAL COSTS

Finally, managers will want a successful screening programme to produce more health gain than the other possible uses for available resources. Toxoplasmosis screening definitely fails on this count. It must cost some money, and as it reduces expected utility, should not be performed. Screening for Down's syndrome and cystic fibrosis add value but do not necessarily add as much as other activities. Screening for Down's syndrome probably provides more utility per pound than most other health interventions because, if we take into account the cost of caring for the Down's syndrome babies whose birth is prevented, it may even save money.¹² If not, we could cost the programmes and turn the decision analysis into a cost-utility analysis. This places all health gains or losses on a universal scale so that the programme under investigation can be compared with any other. In our examples we used "disutility of Down's syndrome" scale and placed all other disutilities on this. The most popular universal utility scale is the quality adjusted life year (QALY). In principle the prevention of a Down's syndrome birth could be expressed in QALYs and then the utility of screening could be compared pound for pound with any other activity.

- Wald NJ, Cuckle HS, Densen JW, Nanchahal K, Royston P, Chard T, *et al.* Maternal serum screening for Down's syndrome in early pregnancy. *BMJ* 1988;297:883-7.
- Thornton JG, Lilford RJ, Johnson N. Decision analysis in medicine. *BMJ* 1992;304:1099-103.
- Management brief; risk and reward. *The Economist* 1989;119-20.
- French S. *Decision theory. An introduction to the mathematics of rationality.* Chichester: Ellis Horwood, 1988.
- Boyle MH, Torrance GW, Sinclair JC, Horwood SP. Economic evaluation of neonatal intensive care of very-low-birth-weight infants. *N Engl J Med* 1983;308:1330-7.
- Johnson N, Sutton J, Thornton JG, Lilford RJ, Johnson VA, Peel KR. Decision analysis for best management of mildly dyskyrotic smear. *Lancet* 1993;342:91-6.
- Neuhauser D, Lewicki AM. What do we gain from the sixth stool guiac? *N Engl J Med* 1975;293:226-8.
- Shapiro MM, Matchar DB, Young MJ. The effectiveness of ovarian cancer screening. A decision analysis model. *Ann Intern Med* 1993;118:838-43.
- Hayek FA. *The constitution of liberty.* Chicago: University of Chicago Press, 1960.
- Raiffa H. *Decision analysis: introductory lectures on choices under uncertainty.* Reading, Massachusetts: Addison-Wesley, 1968:271.
- Thornton JG, Lilford RJ, Newcombe RG. Tables for estimation of individual risks of fetal neural tube and ventral wall defects, incorporating prior probability, maternal serum alphafetoprotein and ultrasound examination results. *Am J Obstet Gynecol* 1991;164:154-60.
- Sheldon TA, Simpson J. Appraisal of a new scheme for prenatal screening for Down's syndrome. *BMJ* 1991;302:1133-6.
- Stray-Pedersen B. Toxoplasmosis in pregnancy. *Baillière's Clin Obstet Gynaecol* 1993;7:107-37.
- Peckham C, Hall S, Patel N, Hudson C, Rudd P, Joynton D, *et al.* *Prenatal screening for toxoplasmosis in the UK.* Report of a working group. London: Royal College of Obstetricians and Gynaecologists, 1992.
- Mennie ME, Gilfillan A, Compton M, Curtis L, Liston WA, Pullen I, *et al.* Prenatal screening for cystic fibrosis. *Lancet* 1992;304:214-6.
- Pauker SP, Pauker SG. The amniocentesis decision: ten years of decision analytic experience. Birth defects; original article series. March of Dimes Birth Defects Foundation, 1987;23:151-69.
- Thornton JG, Lilford RJ. Prenatal diagnosis of Down's syndrome; a method for measuring the consistency of women's decisions. *Med Decis Making* 1990;10:288-93.
- Eddy DM. What's going on in Oregon? *JAMA* 1991;266:417-20.
- Marteau TM. Psychological costs of screening. *BMJ* 1989;300:1527.