

Cost-effectiveness analysis: is it ethical?

Alan Williams *Centre for Health Economics, University of York*

Author's abstract

Many clinicians believe that allowing costs to influence clinical decisions is unethical. They are mistaken in this belief, because it cannot be ethical to ignore the adverse consequences upon others of the decisions you make, which is what 'costs' represent. There are, however, some important ethical issues in deciding what costs to count, and how to count them. But these dilemmas are equally strong with respect to what benefits to count and how to count them, some of which expose ethically untenable assumptions about such widely-used clinical criteria as survival rates. One of the advantages of systematic cost-effectiveness analysis is that it exposes these hidden assumptions, and requires explicit judgements to be made about which ethical position is appropriate in a particular policy context. This should have the important incidental benefit of improving the accountability of policy-makers to the community they are serving.

Introduction

1. In this paper I want to tackle two kinds of ethical issue that have been raised by critics of the cost-effectiveness approach. The first is whether cost-effectiveness analysis *per se* is unethical when applied to medical care. The second concentrates on particular assumptions that are usually made within cost-effectiveness studies. I will deal with each issue in turn.

Is cost-effectiveness analysis unethical *per se*?

2. Many clinicians disapprove of the introduction of economic considerations into priority-setting in medicine, believing that letting costs influence clinical decisions or policies is simply unethical. A particularly sharply worded protest along these lines appeared in the *New England Journal of Medicine* in 1980. Here is an extract:

'Of late an increasing number of papers in this and other journals have been concerned with the "cost-effectiveness" of diagnostic and therapeutic procedures. Inherent in these articles is the view that

choices will be predicated not only on the basis of strictly clinical considerations but also on economic considerations It is my contention that such considerations are not germane to ethical medical practice A physician who changes his or her way of practising medicine because of cost rather than purely medical considerations has indeed embarked on the "slippery slope" of compromised ethics and waffled priorities' (1).

That is a tough act to follow and I am not proposing to chase all the hares that have been set running. I will concentrate on the central issue of whether it is ethical for doctors (and presumably all others involved in patient care) to take costs into account when choosing a course of action.

3. The key to resolving this conflict is to understand the fundamental meaning of the economists' notion of cost. To an economist 'what will it cost?' means 'what will have to be sacrificed?', and this may be very different from 'how much money will we have to part with?'. So if someone says to me that they must have something *no matter what it costs*, I take them to mean that they must have it *no matter what sacrifices have to be made*. And it is always easier to make such statements if the costs (or sacrifices) are going to be borne by somebody else!

4. Transferring that little homily back into the field of medical practice, anyone who says that no account should be paid to costs is really saying that no account should be paid to the sacrifices imposed on others. I cannot see on what *ethical* grounds you can ignore the adverse consequences of your actions on other people. You can do so on bureaucratic or legalistic grounds, of course, by saying 'they are not my responsibility', but we all know into what an ethical morass that line of defence leads. The word we normally use to describe people who behave without regard to the costs of their actions is not 'ethical' but '*fanatical*', and I think that fanaticism is just as dangerous in medicine as it is in other walks of life. So I conclude that a caring, responsible and ethical doctor *has* to take costs into account. Indeed, it is *unethical not* to do so!

5. There is, however, also a somewhat more subtle approach which essentially argues that it is clinical freedom that is being challenged by those pursuing the

Key words

Cost-effectiveness; ethics; resource allocation; costs; QALY; benefit measurement.

cost-effectiveness approach, and whereas clinical freedom has a strong moral base in traditional medical ethics, economics has no such moral base, and must therefore be treated with the utmost suspicion. According to the Stanford University Medical Center Committee on Ethics, the six basic principles of medical ethics are:

1. Preserve life
2. Alleviate suffering
3. Do no harm
4. Tell the truth
5. Respect the patient's autonomy
6. Deal justly with patients

That committee acknowledged that these principles frequently come into conflict with each other, and went on to observe that: 'resolving such conflicts is central to the art of medicine' (2).

6. We health economists have no difficulty in accepting that those same six principles should guide our professional activity, but it may well be that we put rather more weight on the last one (about dealing justly with patients) than doctors have in the past. Again it comes back to how one takes into account the sacrifices borne by others when deciding what 'dealing justly with patients' actually entails. It clearly requires an appeal to some underlying theory of distributive justice, and in practice this is usually some egalitarian principle, though typically so imprecisely formulated as not to constitute a very clear guide to action. But since this imprecision is a notable characteristic of all the other five principles too, it simply leaves us where we were before, except that now it is those involved in *the art of health service management* who are resolving the conflicts, rather than the doctors. And since issues of community-wide 'just dealing' between patients will go beyond the scope of any one doctor's realm of action, it could be argued that if the judgements made by a particular doctor (exercising his clinical freedom) clash with those of someone with authority from the community to allocate scarce resources across rival claimants, the clinical freedom of the doctor has the weaker moral claim, and can legitimately be constrained accordingly.

Are particular assumptions unethical?

The measurement of resource consequences

7. If we now move from the general to the particular, it will be useful to consider separately the measurement of resource consequences and the measurement of health consequences. This section will therefore concentrate on the former, and the next section on the latter.

8. It may come as a surprise that the measurement of resource consequences (for example costs) raises any issues whatever beyond checking for accuracy the boring arithmetic of boring arithmetic people (which can surely be left to other boring arithmetic people) ... by which is clearly meant accountants, not economists!

But there are in fact two big ethical issues to consider. Firstly, should we include amongst the benefits of a treatment any reductions in the indirect costs of illness that are due to that treatment? Secondly, should we ignore the distribution of the costs of the health-care system between different groups in the community?

9. But before tackling either of those questions, I must make clear the context within which these questions get posed. In a health-care system guided by market forces (for example willingness-and-ability-to-pay on the demand side, and profit-seeking on the supply side) there need be no ideological unease about who gets the benefits or who bears the costs, provided that the distribution of purchasing power, and the distribution of market power, are both considered ethically acceptable. No health-care system I know of works wholly on that basis, however, and many, including the British National Health Service (NHS), have explicitly rejected it in principle (though it persists to a limited extent in practice). An egalitarian stance pervades these alternative systems, though as we shall see shortly this egalitarian stance is not sufficiently well specified to offer clear guidance to analysts as to what the distributional policy of the system actually is. And within this large set of predominantly (but vaguely) egalitarian systems, are a smaller number in which it has been decided that the best way to finance them is by taxation (and usually by taxes levied centrally by the national government). In what follows I shall assume that this is the kind of system we are considering.

10. So let us return to the first issue, which was whether the reduction of so-called 'indirect costs' should be counted amongst the benefits of a treatment in a cost-effectiveness analysis (ignoring here any technical difficulties that might arise in trying to do so). The typical situation is as follows: earlier return to work, or less time spent off work, are cited amongst the benefits of a treatment. But we can interpret this phenomenon in several different (but not mutually exclusive) ways. First of all it may be taken as a proxy for improved health status (for example reduced disability). Secondly it may be regarded as the fulfilment of a satisfying social role having value in itself (renewed contact with workmates, a more interesting life, greater self-esteem from doing something useful and valued by others, etc). Thirdly, it may be significant as a source of income (and hence of better living standards) for the whole household. Finally, it may increase national output and thereby benefit the community generally. It is in this last respect that the ethical problem arises.

11. It is sometimes claimed that a particular treatment 'pays for itself' because the increase in national output (decrease in indirect costs) that it brings about is larger than the service costs entailed (output used) in providing the treatment. This increase in output is usually measured by the change in the gross earnings of the treated patients (reflecting what their extra input is worth to their employers, and

ultimately what the consequently increased output is worth to the consumers). A treatment which benefits the unemployable or the retired segments of the population generates no such additional benefits, nor will any such benefits show up for those whose (unremunerated) work is within the home. Even within the working population, this benefit will be greater for the highly paid than for the lowly paid. So in all these respects the counting of indirect costs seems to run counter to the principle that the provision of treatment should not be influenced by whether you are rich or poor.

12. But we are now on the horns of a dilemma, for ignoring such changes in national output is not the same as denying their existence. If the result of ignoring such changes is to concentrate treatments upon the non-working population, the resources available to provide health care and many other good things in life will be less than they might have been (average real income per head will be lower), and it is well known that health is strongly correlated with real income.

13. So far the British health economists have got themselves off this hook by a rather neat argument. Observing that the British economy is currently operating with a very high level of unemployment, it is pointed out that if an otherwise employed person is 'off sick', there are plenty of people willing to fill the void, so all that will happen is that there will be some small 'frictional' losses while the system readjusts, but at the end of the day each sick person 'off work' will have been replaced by some otherwise unemployed person now 'in work', and gross national product will be virtually unchanged. So it is concluded that there are good *economic* arguments for ignoring these indirect costs in the British context. Until such time as full employment returns we need not face the ethical question directly. But what do people who come from countries which do enjoy near-full employment propose to do?

14. The second ethical dilemma on the resource side concerns the distribution of the costs of running the health service. At present we ignore this issue completely. We justify this by arguing that, because the NHS is not financed by earmarked taxes, we cannot identify the taxes which are higher because of the NHS, so it is quite impossible to answer the question 'who bears the costs of the health service?', so there is no point in asking it!

15. But there are increasingly frequent instances where someone advocates the introduction of some additional levy specifically earmarked for the NHS, because it is well known that the NHS is the one public service for which the majority of the population would be prepared to pay more taxes. The question then arises as to whether we care about the incidence of this *extra* taxation. It would be argued that since we are committed to ignoring the distribution of benefits between the rich and the poor, we should ignore the distribution of costs in the same way. Or we could

imbibe the pure milk of the Communist manifesto and say that the guiding principle is 'from each according to his ability, to each according to his need', where 'ability' here refers to ability to pay. The classic alternative taxation principle is that taxes should be proportional to benefits. This is an issue which could become quite hot in the near future, and which could greatly complicate the conduct of cost-effectiveness studies, whichever way we turn.

Are particular assumptions unethical? The measurement of benefits

16. It is in the area of benefit measurement that I have encountered the most intense ethical objections to the cost-effectiveness approach. They vary from quite sweeping denials of the right of anyone to sit in judgement on the value of another person's life, to more specific accusations of ageism, racism, sexism, etc. The criticisms seem to reach a particularly high pitch of excitement when the quality-adjusted-life-year (or QALY) is used as the measure of effectiveness in cost-effectiveness studies.

17. I will not spend much time here on the alleged immorality of one person making judgements about the value of another person's life (or, more correctly, on the value of improvements in another person's health). I think such judgements are inescapable in a system which is expected to behave in a non-capricious manner in discriminating between the well and the ill, between the severely ill and the slightly ill, and between those likely to benefit from a particular treatment and those unlikely to do so, in order that some systematic priority-setting can take place in the face of inescapable resource constraints. The supposedly more ethical alternative of making these decisions by lottery certainly has the advantage of irresponsibility (if indeed that is an advantage), but seems to me quite inhuman and uncaring, and most people I have spoken to about it find it quite unacceptable. But I think it has a serious internal contradiction which flaws it fatally, which is that lotteries do not spring fully formed from Heaven. They are invented by people. These people have to decide who is eligible to enter this lottery, what the prizes are, how soon and how often you can re-enter the lottery if you fail to win the first time, whether 'tickets' (especially winning tickets) can be traded or given away, and so on. It seems to me to be the *beginning* of a new discussion about discrimination, which merely takes the place of the old one, but does not get us off that particular ethical hook. Having accepted the inescapability of such judgements, what is of more practical interest is an examination of the ethical implications of those judgements.

18. The first general point that has to be made is that *every* effectiveness measure implies some value judgement. These frequently go unrecognised, because the effectiveness measure has come to be so widely used that it is conventionally accepted as the appropriate *technical* way of doing things.

Take for instance the two-year survival rate as a criterion for choosing between rival treatments. It carries the following implications:

- (1) to survive less than two years is of no value
- (2) having survived two years, further survival is of no additional value
- (3) it does not matter with what *quality* of life people survive to two years
- (4) it does not matter who you are

The only one of these implicit assumptions that is acceptable to me is the last one, but as I propose to indicate shortly, even that is not acceptable to everybody. So on ethical grounds the two-year survival rate may well prove to be a totally unacceptable measure of effectiveness! It's a bit worrying, isn't it?

19. The objection that survival rates pay no regard to quality of life can be overcome by adopting the QALY as the effectiveness measure, and then investigating quite explicitly the extent to which people are prepared to sacrifice quality-of-life to increase life-expectancy, or vice-versa. Since people are likely to have different views on this important issue, we have to decide how the different valuations are to be brought together when making collective decisions. A typical technical solution is to take the arithmetic mean, which in the case of a skewed distribution gives a lot of weight to the upper extreme. This can be avoided by using median or modal values, or some more complex mean (for example the geometric), but what we are in fact doing here is deciding whose values shall count for how much in whatever policy issue is being addressed by the study (and this is equally true of clinical trials as for cost-effectiveness studies). This has quite important implications for the ethical problem of 'dealing justly with patients', so maybe that boring arithmetic is not as unimportant as it looks!

20. I noted earlier the strongly-held, but vaguely-articulated, egalitarian notions which supposedly guide many health-care systems. When aggregating benefits across people we have to know what it is that is supposed to be equally valued across people, otherwise aggregation is impossible. So the analysts are having to fill the void (wittingly or unwittingly) by adopting a precise stance on matters on which society is not offering clear guidance. There are many potential targets for such an egalitarian ethic. Even if attention is restricted to outcome measures (and some people argue that process measures are equally relevant) the list of candidates is quite long. A few obvious ones are a life-year, a QALY, whether the person has already had 'a fair innings', the 'rest of your life' (favoured by some philosophers because it minimises the scope for judgements about length and quality), or some other differentiating characteristic to do with who is getting the benefit. Like all other analysts, in order to get the work done I have filled that void in my own field, by explicitly assuming that a QALY is a QALY is a QALY

no matter who gets it. But I do not feel easy about this, so I have been trying to find out what the general public thinks about these matters, and what those responsible for priority-setting in health care think about these things. The difficulty with such empirical work is, of course, that most of the time they do not think about these things at all, so it is quite difficult to elicit their views in a systematic and reliable way. And although I do not pretend to have cracked this problem yet, let me offer you some preliminary results to think over.

21. Whether a unit of benefit should have a different value depending on who will get it pervades much of the discussion about distributive ethics. The characteristics that are frequently mentioned in these debates are age, sex, marital status, whether with or without children, occupation, and whether the person has cared for his or her own health. A few years ago we conducted a survey amongst nearly 400 randomly-selected adult citizens of York asking them at what stages in life they considered it most important to be healthy. The life stages considered were:

As infants	When at peak of earning power
When starting school	When looking after elderly relatives
When starting work	When just having retired from work
When setting up home for the first time	When coping with death of spouse
When bringing up children	When getting very old

The results surprised me. There was a very clear consensus that the most important time to be healthy was when bringing up children, which is a time in people's lives when they are in fact usually quite healthy, but which is obviously also a time when people feel extremely vulnerable if they are not healthy. Running a close second was when you are an infant, the reason usually given here being that a healthy start in life is a good investment for the future. All other life stages were far behind these two in importance.

22. I later tried out a somewhat different approach, asking a convenience sample of some 80 people on what grounds they thought the NHS should discriminate between different sorts of people when determining priorities. The largest single group (40 per cent of respondents) thought there should be no discrimination whatever. In the other 60 per cent the preferred bases of discrimination were according to whether people had or had not cared for their own health (which was particularly prominent amongst doctors and health-service managers), next in importance being a preference for the young. Roughly equal, in third place, were those looking after children (except amongst health-service managers!), and those looking after elderly relatives (especially amongst secretaries, who were mostly middle-aged women!)

23. A much larger survey has been carried out in Cardiff on just over 700 people selected randomly from the electoral register. They were asked to choose between people with different characteristics when treatment could only be given to one of them. The results indicated a strong preference for the young over the old, except for the very young, where an eight-year-old would be given precedence over a two-year-old. Incidentally, the older respondents also manifested this preference for the young over the old, typically adducing by way of justification the 'fair innings' principle. In addition, married people were preferred to single people, non-smokers to smokers, and light drinkers to heavy drinkers. No clear view emerged on the other dimensions tested (sex and occupation).

24. I later did a similar survey on the senior members of an English health authority with strikingly similar results. Again the young were to be given precedence over the old, non-smokers over smokers and light drinkers over heavy drinkers. We tested in addition the dimension of having children of school age versus having no children, eliciting a very strong preference for the former. Again sex and occupation made no difference. These data have already convinced me that there is a very strong consensus in Britain concerning discrimination by age, by whether someone has young children, and by smoking and drinking habits. It probably influences treatment priorities almost unconsciously at local levels. But should *we* be taking all this into account in clinical trials and in cost-effectiveness studies? And if so, how are we to do so? At present I am sticking uneasily to a QALY is a QALY is a QALY, pending further clarification of just *how much* extra weight is to be given to the favoured categories over the unfavoured ones.

Conclusions

25. I hope that I have now convinced you not only that

the evaluation of health-care activities is an ethical minefield, strewn with explosive material not easily detected by the naked eye, but also that bringing this material out into the open and analysing it (both by logical discourse and by empirical enquiry) is an important extension of the analyst's role. I think it is our duty to rush in where others fear to tread, even if in the process we find ourselves being maligned as insensitive troublemakers, and even if the misguided criticize our analytical techniques because they require quite strong ethical assumptions to be made. It is not that any of these analytical techniques are ethical or unethical *per se*, it is more a matter of ensuring that their particular ethical assumptions are appropriate in the context in which they are being used. To do that requires *us* to be clear about the ethical assumptions built into our studies, but it also requires our 'clients' to be clear about what ethical assumptions are appropriate in *their* worlds.

I suspect that we will make faster progress with our task than they will with theirs, but perhaps it is precisely through our questioning that their position will become clearer, both to themselves and to everyone else. And with increasingly insistent demands for greater professional and political accountability in the provision of health, that must be A Good Thing!

Alan Williams is Professor of Economics at the Centre for Health Economics, University of York.

References

- (1) Loewy E L [letter]. Cost should not be a factor in medical care. *New England journal of medicine* 1980; 302: 697.
- (2) Ruark J E, Raffin T A and the Stanford University Medical Center Committee on Ethics. Initiating and withdrawing life support. *New England journal of medicine* 1988; 318: 25-30.