Serie Occasional notes on economics - *British Medical Journal*

1. Raftery J *Economic evaluation: an introduction* *BMJ* 1998;316:1013-1014
2. Byford S, Raftery J *Perspectives in economic evaluation* *BMJ* 1998;316:1529-1530
3. Palmer S, Torgerson DJ *Definitions of efficiency* *BMJ* 1999;318:1136-1136
4. Palmer S, Byford S, Raftery J *Types of economic evaluation* *BMJ* 1999;318:1349-1349
5. Torgerson DJ, Raftery J *Measuring outcomes in economic evaluations* *BMJ* 1999;318:1413-1413
6. Palmer S *Opportunity cost* *BMJ* 1999;318:1551-1552
7. Torgerson DJ, Raftery J *Discounting* *BMJ* 1999;319:914-915
8. Byford S, Raftery J, Torgerson DJ *Cost of illness studies* *BMJ* 2000;320:1335
9. Raftery J *Costing in economic evaluation* *BMJ* 2000;320:1597

Notes edited by James Raftery (J.P.RAFTERY@bham.ac.uk)
1. BMJ 1998;316:1013-1014 (28 March)

Education and debate - *Economics notes*

**Economic evaluation: an introduction**

James Raftery, *director.*

Health Economics Facility, Health Services Management Centre, University of Birmingham, Birmingham B15 2RT

Economic evaluation has increasingly become part of modern health care. Rising costs, often associated with new technologies, and spending limits have prompted a search for greater efficiency. This need to demonstrate the relative cost effectiveness of new health technologies has led some countries, specifically Australia, to make economic evaluation a requirement for public sector funding of new drugs. Furthermore, the American state of Oregon used economic evaluation in defining what services should be included in Medicare (although the rationing eventually implemented relied mainly on effectiveness rather than cost effectiveness). This growing requirement to demonstrate the efficiency of new technologies means that economic evaluation is increasingly specified in research grants from both the NHS and pharmaceutical companies.

Economic theory, which takes private markets and rational individual decision making as the norm, has developed techniques—primarily cost benefit analysis—to evaluate programmes funded by the public sector. As the earliest forms of cost benefit analysis measured both costs and benefits in monetary terms, the term cost benefit analysis has come to mean those analyses which measure outcomes in monetary terms. Other forms, specifically cost effectiveness and cost utility analysis, have been developed to cover analyses in which outcomes are measured in health related terms. The results of such studies are usually ratios of costs to outcome. The most generalisable, cost per quality adjusted life year (QALY) gained, has provoked controversy and, despite its apparent simplicity, raises many technical complexities.

Economic theory favours measuring costs and benefits in monetary terms because it avoids the problems of measuring and valuing non-monetary benefits, such as health gain or patient satisfaction. The branch of economics that deals with individuals—welfare economics—uses great ingenuity to avoid measuring the "utility" or satisfaction of different individuals. According to welfare economics, rational individuals will maximise their utilities and that of society in perfectly competitive markets. However, as discussed in this series, most healthcare evaluations have to grapple with benefits that are specific to health care, and about which consumers often have limited information. This pushes conventional economics thinking to its limits, raising difficult welfare comparisons between individuals and over time, leading some health economists to propose an "extra welfarist economics."

Although many problems of economic evaluation relate to measuring and valuing benefits, some also relate to costs. Economics defines costs much more broadly than accountancy. The concept of "opportunity cost" defines the cost in terms of the next best opportunity foregone. As welfare economics takes a societal perspective, the relevant opportunity cost is that to society rather than to an individual or an organisation. Opportunity cost includes not only the direct costs of treatments, but also the knock on costs of treatments averted or postponed and the costs to patients such as time spent waiting or off work or due to being cared for. Such definitions of costs, while comprehensive, are rarely available from routine sources.

**Terminology**

*Welfare economics* is a branch of economics concerned with maximising social welfare. It assumes rational individuals who maximise their utilities, and that the overall welfare of society is a function of individual utilities. *Health economics as welfare economics* applies
welfare economics to health care. *Health economics as extra-welfarist* is concerned with maximising health which may include both individual and social preferences. It builds on but goes beyond the individualist focus in welfare economics.

As the popularity of economic evaluation of health care has increased, so too has the demand for rigour in its methods. While some have argued long and persuasively for measures such as cost per QALY, others have pointed to the limited range of interventions that have been evaluated and to the lack of standard methods in deriving such estimates as are available. In response, standardised methods for economic evaluation have been suggested. The problems of combining costs and benefits in evaluating health care have led some to caution against doing economic evaluation as part of clinical trials. The argument relates partly to the difficulty of capturing the full extent of costs in trials, the fact that the power of trials is usually set in terms of benefits not costs, and the fact that trials may be atypical. Modelling and simulations, which have been proposed as alternatives to economic evaluations alongside clinical trials, have, however, also been criticised for being open to bias. Economic evaluation has been dubbed a "half way technology" because of the lack of standardised approaches which requires each study to start anew rather than build on previous work. Others have doubted the benefits of standardisation, favouring instead research on unresolved topics such as outcome measurement, discounting, and the uses to which economic evaluations have been put.

Owing to the increasing importance of using economics in healthcare decision making the *BMJ* will publish a series of economics notes. These do not attempt a comprehensive review of economic evaluation: rather they aim to discuss issues which have arisen in the course of designing and carrying out evaluations. Furthermore, the series will try to clarify economic terminologies.

References


Perspectives in economic evaluation

Sarah Byford, research fellow, a James Raftery, professor of health economics. b

a Centre for Health Economics, University of York, York YO1 5DD, b Health Economics Facility, Health Services Management Centre, University of Birmingham, Birmingham B15 2RT

Before an economic evaluation begins, the perspective of the study should be determined, as it may have implications for trial design. Since economic evaluations are often used to assess the relative efficiency of alternative healthcare interventions, the perspective commonly taken is that of the health service. Because of its foundations in welfare economics, however, health economics is concerned with society's welfare. It therefore argues that an economic evaluation should include the impact of an intervention on the welfare of the whole of society, not just on the individuals or organisations directly involved.

Central to economic theory is the question of how to get the most benefit from the scarce resources available to a society. An economic evaluation which confined itself to the NHS's perspective could determine the mix of interventions that would maximise health outcomes within the limited NHS budget. However, this would not necessarily maximise the welfare of society within resources available (gross national product) for two main reasons.

Firstly, sectors other than the health service may incur costs or benefits as a result of healthcare interventions. Consider for example, the reduction of psychiatric hospital beds, which might seem cost effective from the perspective of the health service but less so from that of society as a whole, including patients' or carers' perspectives. A societal perspective helps detect cost shifting between sectors.

Secondly, a narrow perspective takes no account of alternative uses for resources outside the healthcare sector, which may yield greater welfare to society. The concept of opportunity cost reflects this broad concern for society's total welfare. Because the total economic output of society is limited, choosing to devote resources to health care necessarily implies forgoing the benefits (or opportunities) of using those resources in an alternative sector, such as education or social services, or indeed not raising them as taxes.

Since the NHS is a universally accessible service, funded by taxation, it can be argued that its decisions should be from the perspective of all those who use it and pay for it—the whole population. Epidemiology and economics share a concern with populations. Adopting a societal or population perspective facilitates policies aimed at maximising the welfare gains to society, or minimising the losses. Excluding particular groups could hinder this process and be less equitable than attempting to include the views of all.

The utilitarianism that underpins welfare economics has been characterised as a poor basis for a personal philosophy (uncaring, calculating, and consequentialist) but an appropriate basis for public policy. While a utilitarian, societal perspective can be justified in principle, its practice can be difficult. Ethical dilemmas can arise between individual and societal perspectives. When an individual is denied (even possibly ineffective) treatment the lack of a societal consensus can become stark. How individual preferences are to be aggregated to a societal preference remains a theoretical and practical challenge.

To certain organisations, a societal perspective may seem unnecessary. A clinical directorate faced with difficult decisions within a tight budget may take a directorate perspective, in turn
requiring the wider organisation to act to prevent cost shifting or other undesirable effects. For-profit organisations, such as pharmaceutical companies, may well take a narrow financial perspective, in turn leading to regulation by the state to safeguard wider concerns. The emphasis in guidelines for pharmacoeconomic evaluations to include a societal perspective, however, is striking.

For policy purposes, study comparability is enhanced by adopting a societal perspective as a norm. A report for the US Public Health Service recommended a societal perspective, for which it saw welfare economics as providing the best theoretical framework, for use in reference case economic evaluations—those that aim to provide results that could be compared throughout the healthcare system.

In practice, it may not always be possible for all the relevant costs and benefits to be included in an economic evaluation because of funding or time constraints. A good case can be made for excluding particular effects if they are likely to have little impact on the overall results. Pretrial literature reviews and modelling can help prioritise items of importance. A "reduced list" method has been shown to capture most relevant costs in mental health service evaluations, with the five most costly services accounting for 94% of the total cost and the next five for only 4%. Such short cuts require further evaluation before they are more widely applied. A similar analysis on costing screening for colorectal cancer found reduced list costing to be less successful. As economic evaluation becomes more standardised it may be possible to justify such limited perspectives for costing particular diseases or services.

At the very least economic evaluations should be explicit about the perspective they adopt. The exclusion of items, whether for practical reasons or as a result of pretrial assessments, must be made explicit, explained, and discussed in terms of their likely influence on the final results. Studies with non-societal perspectives may result in suboptimal resource allocation decisions and a corresponding loss in the total welfare of society.


Decision makers are increasingly faced with the challenge of reconciling growing demand for health care services with available funds. Economists argue that the achievement of (greater) efficiency from scarce resources should be a major criterion for priority setting. This note examines three concepts of efficiency: technical, productive, and allocative.

Efficiency measures whether healthcare resources are being used to get the best value for money. Health care can be seen an intermediate product, in the sense of being a means to the end of improved health. Efficiency is concerned with the relation between resource inputs (costs, in the form of labour, capital, or equipment) and either intermediate outputs (numbers treated, waiting time, etc) or final health outcomes (lives saved, life years gained, quality adjusted life years (QALYs)). Although many evaluations use intermediate outputs as a measure of effectiveness, this can lead to suboptimal recommendations. Ideally economic evaluations should focus on final health outcomes.

Adopting the criterion of economic efficiency implies that society makes choices which maximise the health outcomes gained from the resources allocated to healthcare. Inefficiency exists when resources could be reallocated in a way which would increase the health outcomes produced.

Technical efficiency refers to the physical relation between resources (capital and labour) and health outcome. A technically efficient position is achieved when the maximum possible improvement in outcome is obtained from a set of resource inputs. An intervention is technically inefficient if the same (or greater) outcome could be produced with less of one type of input. Consider treatment of osteoporosis using alendronate. A recent randomised trial showed that a 10 mg daily dose was as effective as a 20 mg dose. The lower dose is technically more efficient.

Productive efficiency cannot, however, directly compare alternative interventions, where one intervention produces the same (or better) health outcome with less (or more) of one resource and more of another. Consider, for example, a policy of changing from maternal age screening to biochemical screening for Down's syndrome. Biochemical screening uses fewer amniocenteses but it requires the use of another resourcebiochemical testing. Since different combinations of inputs are being used, the choice between interventions is based on the relative costs of these different inputs. The concept of productive efficiency refers to the maximisation of health outcome for a given cost, or the minimisation of cost for a given outcome. If the sum of the costs of the new biochemical screening programme is smaller than or the same as the maternal age programme and outcomes are equal or better, then the biochemical programme is productively efficient in relation to the maternal age programme. In health care, productive efficiency enables assessment of the relative value for money of interventions with directly comparable outcomes. It cannot address the impact of reallocating resources at a broader levelfor example, from geriatric care to mental illnessbecause the health outcomes are incommensurate.
Allocative efficiency To inform resource allocation decisions in this broader context a global measure of efficiency is required. The concept of allocative efficiency takes account not only of the productive efficiency with which healthcare resources are used to produce health outcomes but also the efficiency with which these outcomes are distributed among the community.\textsuperscript{6} Such a societal perspective is rooted in welfare economics and has implications for the definition of opportunity costs. In theory, the efficient pattern of resource use is such that any alternative pattern makes at least one person worse off. In practice, strict adherence to this criterion has proved impossible. Further, this criterion would eliminate as inefficient changes that resulted in many people becoming much better off at the expense of a few being made slightly worse off. Consequently, the following decision rule has been adapted: allocative efficiency is achieved when resources are allocated so as to maximise the welfare of the community.\textsuperscript{6}

Thus technical efficiency addresses the issue of using given resources to maximum advantage; productive efficiency of choosing different combinations of resources to achieve the maximum health benefit for a given cost; and allocative efficiency of achieving the right mixture of healthcare programmes to maximise the health of society. Although productive efficiency implies technical efficiency and allocative efficiency implies productive efficiency, none of the converse implications necessarily hold. Faced with limited resources, the concept of productive efficiency will eliminate as "inefficient" some technically efficient resource input combinations, and the concept of allocative efficiency will eliminate some productively efficient resource allocations.

Footnotes

These notes are edited by James Raftery J.P.RAFERY@bham.ac.uk

References


Types of economic evaluation

Stephen Palmer, research fellow a, Sarah Byford, research fellow a, James Raftery, professor b.

a Centre for Health Economics, University of York, York YO1 5DD, b Health Economics Facility, Health Services Management Centre, University of Birmingham, Birmingham B15 2RT

The pursuit of efficiency in the healthcare sector requires priority to be given to those treatments which provide the greatest benefit per unit of cost. Alternative interventions often have to be compared to determine whether a change in the mix of interventions would increase efficiency. Although economic evaluations approach costs in a common format, they differ in the way they approach benefits. These differences play a critical role in developing criteria for efficiency.1

Cost benefit analysis involves measuring costs and benefits in commensurate terms, usually monetary. Welfare economics shows that under certain conditions any net excess of monetary benefits over costs represents the gain in welfare by society.1 Cost benefit analysis makes it possible to determine, firstly, whether an individual intervention offers an overall net welfare gain and, secondly, how the welfare gain from that intervention compares with that from alternative interventions. Increased use of interventions with the greatest net gain will increase efficiency. By valuing all costs and benefits in the same units, cost benefit analysis compares diverse interventions using the net benefit criterion. Cost benefit analysis thus simultaneously addresses issues of productive and allocative efficiency.

Practical measurement difficulties and objections to valuing health benefits in monetary terms have limited the use of cost benefit analysis in health care, though recent approaches using the concept of “willingness to pay”2 have revived interest in it.

Cost effectiveness analysis measures health benefits in natural units such as life years saved or improvements in functional status (units of blood pressure or cholesterol). Since costs and benefits are measured in non-comparable units, their ratio provides a yardstick with which to assess relative (productive) efficiency.3 This decision rule does not, however, enable us to evaluate the relative efficiency of interventions which provide more benefit at greater cost or less benefit at lower cost.4

If an intervention is both more expensive and more effective than an alternative, then the criterion for efficiency becomes the ratio of the net increase in costs to the net increase in effectiveness (the incremental cost effectiveness ratio). However, the additional expense of the new intervention means that resources have to be redirected from elsewhere. An economic evaluation assesses whether or not the additional benefits generated by the new intervention are greater than the loss in benefits from the reduction in other programmes that is, is the reallocation efficient? A major limitation of cost effectiveness analysis is its inability to compare interventions with differing natural effects.5 For example, interventions aimed at increasing life years gained cannot be directly compared with those which improve physical functioning. Cost effectiveness analysis therefore cannot directly address allocative efficiency.4

Cost utility analysis is an adaptation of cost effectiveness analysis which measures an intervention’s effect on both the quantitative and qualitative aspects of health (morbidity and mortality) using a utility based measure such as quality adjusted life years (QALYs).6 Like cost
effectiveness analysis, relative efficiency is assessed using an incremental ratio, here a cost utility ratio. An intervention is deemed productively efficient, relative to an alternative, if it results in higher (or equal) benefits at lower cost. The use of a single measure of health benefit enables diverse healthcare interventions to be compared so cost utility analysis can address both productive efficiency and allocative efficiency.

In cost utility analysis the optimal decision rule involves ranking the incremental cost utility ratios of different interventions and selecting those with the lowest ratios (best value) until the budget is depleted. 7 8 The lower the incremental ratio, the higher the priority in terms of maximising health benefits derived from a given level of expenditure.8 The point at which resources are exhausted defines a maximum price for a unit of effectiveness for example, £20 000 per QALY that is affordable within the budget. Eliminating interventions with an incremental cost above this price in favour of those with lower incremental costs would be considered an improvement in allocative efficiency.

In practice huge difficulties exist in obtaining enough information to facilitate such ranking. Information is required on the full costs and benefits of all health problems and all alternative interventions. Within each health problem subgroups with different levels of potential health gain must be distinguished.

Footnotes

These notes are edited by James Raftery (J.P.RAFTERY@bham.ac.uk)

References
Measuring outcomes in economic evaluations

David Torgerson, research fellow a, James Raftery, professor b.

a National Primary Care Research and Development Centre, Centre for Health Economics, University of York, York YO1 5DD, b Health Economics Facility, Health Services Management Centre, University of Birmingham, Birmingham B15 2RT

To make judgments about efficiency economic evaluation of health care has to compare health outcomes, however measured, with costs. Three main approaches exist to measuring outcomes: clinical end points, quality of life measures, and willingness to pay.

The simplest outcome measure to use in a trial is a clinical one, such as a reduction in the number of strokes or changes in blood pressure. Health economists use such measures to construct cost effectiveness ratios.1 For example, in a trial aimed at preventing hip fractures a cost effectiveness ratio might be cost per averted hip fracture.

Measuring outcome in terms of clinical endpoints has the disadvantage that comparisons between different healthcare treatments are difficult. This is only partly solved when trial endpoints include mortality. Although estimates of cost per life gained or life year gained allow comparisons between very different therapies, using survival as an outcome measure for an economic evaluation is problematic. Firstly, few clinical trials are powered to detect mortality differences. Secondly, many treatments affect morbidity rather than mortality. Thirdly, even when survival is an appropriate end point, reductions in mortality may be at the expense of reductions in quality of life.

Measures of quality of life which go beyond both clinical and mortality end points are becoming more common. Quality of life measures may be condition specific, generic, or utility based. Condition specific measures comprise questions about particular symptoms which treatment aims to resolve. For example, the Roland and Morris back pain scale asks about a patient's back pain and how this limits functional activity.2 In contrast, a generic measure such as the SF36 asks questions about an individual's general health. Finally, utility measures, such as the EuroQol, go beyond generic measures: they have interval/ratio properties and are preference based.

By being based on scales with interval properties, utility measures enable different interventions to be compared. If health intervention A improved patients' health, on average, by 10 points on a utility scale and intervention B by 5 points, then intervention A is twice as effective. Most generic quality of life measures lack interval properties.

Furthermore, the valuation of utility measures is based on societal preferences. Although many condition specific measures are based on patients' valuations, those used in generic and utility based measures tend to use population valuations.

Utility measures tend to be relatively insensitive to important changes in health status. Unless sample sizes are extremely large, reliance on utility measures alone runs the risk of type II errors concluding that there is no important quality of life gain when there is. Health economists often recommend using utility measures alongside other, more sensitive, measures of outcomes.

Sometimes the benefits of healthcare interventions go beyond clinical or quality of life changes. Most couples undergoing in vitro fertilisation will not have a baby: for those, coming to terms with
their infertility may be a benefit. Similar considerations apply to non-health benefits, such as respect of patients' autonomy and dignity. To measure these benefits, the techniques of willingness to pay and conjoint analysis have been suggested.

In willingness to pay, the outcome of a healthcare procedure and its alternative(s) are described and patients asked how much they would be willing to pay. Procedures with the highest values are preferred. Besides capturing non-health dimensions, this technique enables benefits to be expressed in monetary terms, allowing cost benefit analysis to take place. Use of willingness to pay in evaluating cystic fibrosis screening showed that benefits other than knowing carrier status were important.

Conjoint analysis presents patients with a list of pairwise choices of a health intervention. For example, whether patients preferred their general practitioner's surgery to have longer opening times combined with a night time deputising service or shorter day time opening combined with the general practitioners doing their own on call. The various attributes of alternatives can be weighted to generate utilities. That option with the best cost utility ratio is the most efficient.

References
The concept of opportunity cost is fundamental to the economist's view of costs. Since resources are scarce relative to needs, the use of resources in one way prevents their use in other ways. The opportunity cost of investing in a healthcare intervention is best measured by the health benefits (life years saved, quality adjusted life years (QALYs) gained) that could have been achieved had the money been spent on the next best alternative intervention or healthcare programme.

Opportunity cost can be assessed directly with cost effectiveness or cost utility studies. When two or more interventions are compared cost utility effectiveness analysis makes the opportunity cost of the alternative uses of resources explicit. Cost effectiveness ratios, that is the £/outcome of different interventions, enable opportunity costs of each intervention to be compared.

Although the concept of opportunity cost is fundamental, incorrect conclusions can result from difficulties in applying the concept. Firstly, the study perspective (societal, patient, etc) is critical since it determines which costs and effects to include in the evaluation. A societal perspective incorporates all the costs and benefits regardless of who incurs or obtains them. More restricted perspectives may mask the fact that costs are simply being shifted to another sector rather than being saved.

Secondly, the choice of comparisons can play a crucial part in cost effectiveness analysis, affecting the measurement of opportunity cost. Ideally an intervention should be compared with all relevant interventions, including doing nothing. Without a "do nothing" baseline, the best of two generally undesirable options may be chosen. Sometimes, however, the do nothing option may be unethical, such as when a new treatment is being compared with one that has been shown to be beneficial. Partly for this reason, many studies compare particular interventions with existing practice which may or may not be well defined. Failure to select an appropriate comparator may make the intervention appear more cost effective than it should, leading to wrong estimates of the opportunity cost.

Thirdly, the incremental rather than average cost effectiveness ratio should be estimated. The average cost per benefit (calculated by dividing the total cost of an intervention by the total benefits) may be less appropriate than the incremental ratio (derived by dividing the additional (incremental) costs by the additional (incremental) benefits). A recent study showed that the incremental cost effectiveness ratio for maternal age screening was 27% higher than the average ratio and concluded that the failure to consider incremental ratios could mislead decision makers about the opportunity cost of screening in Down's syndrome.
Resources used in economic evaluations should be valued at opportunity cost, but doing this is difficult (especially in health care, where there is no perfect market), so unit costs tend to be used instead, based on the costs of the various inputs.

Accounting practices do not aim to measure opportunity costs. Opportunity costing generally requires comprehensive, disaggregated data at the individual patient level. Even then, the allocation of overhead and fixed costs is difficult since the cause and effect relation between resources and different users is difficult to determine. Since many economic evaluations use accountancy cost data, the results should be treated with some caution. The prices of pharmaceutical products may be poor estimates of their opportunity cost because the retail price reflects the patent, the regulation of profits by governments, and the sunk research and development of both successful and unsuccessful products. In practice, very few studies attempt to estimate the opportunity costs of drugs, relying instead on prices.

Finally, valuation of resources for which no market exists, such as informal care, or patient time costs, requires methods to derive what economists call "shadow prices" the true social value (or opportunity cost) of non-marketed resources, such as time and informal care.

Health economists disagree about the most appropriate technique for measuring the opportunity cost of time. The best valuation of the opportunity cost of time for working age adults is the wage they are, or could be making, in paid work, varying according to whether the time lost involves lost work or leisure time and the likelihood of being unemployed.

If resources are to be allocated efficiently, then the value of using these resources in alternative ways needs to be made explicit. Despite the importance of this concept, the complexities of its application mean that few studies are even completely explicit about their estimates of opportunity costs. Greater clarity about the perspective of the study could help in clarifying the range of opportunity costs included.

**Term Definition**

- **Opportunity cost**: Benefits forgone by particular use of resources
- **Cost effectiveness ratio**: Ratio of costs to outcomes
- **Incremental cost effectiveness ratio**: Ratio of change in costs to change in outcome

**References**

Education and debate - Economics notes

Discounting

David J Torgerson, senior research fellow a, James Raftery, professor b.

a Department of Health Studies and Centre for Health Economics, University of York, York YO1 5DD, b Health Economics Facility, University of Birmingham, Birmingham B15 2RT

Until recently it has been common practice in economic evaluations to "discount" both future costs and benefits, but recently discounting benefits has become controversial. Discounting makes current costs and benefits worth more than those occurring in the future because there is an opportunity cost to spending money now and there is desire to enjoy benefits now rather than in the future. The reason why current spending incurs an opportunity cost relative to delayed spending is that a monetary investment yields a real rate of return and therefore there is a cost to spending money in the present.

For example, if £100 were invested with a nominal return of 10%, in one year's time it would be worth £110; if inflation was 4% this would result in a real return of £6 on every £100 invested. If for some reason £100 of healthcare spending were delayed for one year then (assuming prudent investment) we could expect that in one year's time we would have £106 for healthcare investment.

To take into account the opportunity cost of investing now rather than waiting one year we have to discount future costs. Therefore, if two healthcare interventions both released £100 in savings but for one we had to wait a year, then, all other things being equal, we would adopt the intervention that saved £100 now. This is because the £100 released now, if invested, would produce an extra £6 in a year's time (with a discount rate of 6%).

Failure to discount the future costs in economic evaluations can give misleading results. For example, an evaluation of cystic fibrosis screening revealed a cost of £80 000 for detecting and terminating one affected pregnancy.1 This cost was compared with the future excess costs of treating an individual with cystic fibrosis, which was estimated to be £5000 a year over 25 years. As cystic fibrosis screening benefits (£125 000) outweighed the costs (£80 000) it was concluded that screening represented good value for money. However, if the averted costs had been discounted (at 6%) then these would have been only £63 917, which alters the study's results (though not if the discount rate were only 4%).

Discounting future costs is uncontroversial and until recently so was the process of discounting health related benefits. The main argument against discounting health benefits is that health, unlike wealth, cannot be invested to produce future gains.2 The Department of Health has thus recommended that health related benefits should not be discounted, 3 4 though more recent advice suggests future health benefits should be discounted but at a very low rate of 1.5%-2%.5

An important reason for discounting future costs and benefits is "time preference," which refers to the desire to enjoy benefits in the present while deferring any negative effects of doing so. Examples of human behaviour which implicitly discount future health effects abound. For instance smoking and drinking give current pleasure while incurring future (discounted) detrimental health effects. Indeed, research has indicated that smokers value future health benefits at a lower rate than non-smokers.6 This desire to enjoy pleasurable benefits in the present time is often reflected in differential pricing of goods and services. Consider the hire of a video for home viewing. Despite
the increased cost of newly released videos, many people are willing to pay the extra cost rather than wait until the price falls.

Failure to discount future health related benefits will tend to show more favourable cost effectiveness ratios compared with discounting. For instance, an evaluation of two view mammography for breast screening showed an undiscounted marginal cost per life year of £1200.7 However, discounting the life years (at 6%) increased the marginal cost per life year by 74%, to £2092.

If future health benefits are not discounted this implies that health gains achieved this year and those achieved in 20 or 30 years are of equal value. As an example, let us assume that about £70 000 is available for hip fracture prevention in 100 women, and there are two strategies under consideration: 10 years of hormone replacement therapy (given to 50 year old women), which prevents 50% of fractures in 30 years' time; or 10 years of calcium and vitamin D (given to 70 year old women), which prevents 30% of hip fractures in 10 years' time.

In the table we show how discounting health benefits alters the relative cost effectiveness of the two interventions to prevent hip fracture. Without discounting, hormone replacement therapy produces a lower cost per avoided hip fracture than vitamin D and calcium. On the other hand if the hip fracture benefits are discounted, then the reverse is true.

In this example the decision on which preventive strategy to adopt is heavily influenced by discounting. Whether we discount health benefits, and at what rate, depends on how much value society places on current health benefits compared with future benefits. Intuitively it would seem best to be able to prevent hip fractures in the next 10 years rather than wait 30 years for this health benefit. What little research there has been into society's preferences about current health benefits compared with future ones suggests that people value current health benefits more highly than future ones and that people actually discount future health gains more highly than future wealth gains.6 Many economists still hold the view that future health gains should be discounted.9 In most economic evaluations the choice of discount rate will not affect the relative ranking of the interventions under evaluation. However, it is good practice to establish whether the evaluation results are critically affected by the discount rate by a sensitivity analysis using different discount rates.

Footnotes

References
Cost of illness studies

Sarah Byford, research fellow a, David J Torgerson, senior research fellow a, James Raftery, professor b.

a Centre for Health Economics, University of York, York YO10 5DD, b Health Economics Facility, Health Services Management Centre, Birmingham B15 2RT

Cost of illness studies are a type of economic study common in the medical literature, particularly in specialist clinical journals. The aim of a cost of illness study is to identify and measure all the costs of a particular disease, including the direct, indirect, and intangible dimensions. The output, expressed in monetary terms, is an estimate of the total burden of a particular disease to society.1

It is widely believed that estimating the total societal cost of an illness is a useful aid to policy decision making, and indeed organisations such as the World Bank and the World Health Organization commonly use such studies.2 However, cost of illness studies have been the cause of much debate among economists. 1 3 4

Two methods of costing illness exist: the prevalence and incidence approaches. The prevalence method is the commonest and estimates the total cost of a disease incurred in a given year. The more data hungry incidence based approach involves calculating the lifetime costs of cases first diagnosed in a particular year, providing a baseline against which new interventions can be evaluated.1

Determining the total cost of an illness is claimed to provide several useful pieces of information. Firstly, it tells us how much society is spending on a particular disease, and by implication the amount that would be saved if the disease were abolished. Secondly, it identifies the different components of cost and the size of the contribution of each sector in society. Such information, it is argued, can help to determine research and funding priorities by highlighting areas where inefficiencies may exist and savings be made. 1 5

There exist, however, several arguments against undertaking and using the results of cost of illness studies. Simply identifying an area of high expenditure does not provide enough information to suggest inefficiency and waste and so should not automatically take precedence for further scrutiny. An inefficient allocation of resources exists when those resources could generate greater benefits if used elsewhere; without an understanding of the benefits (or health outcomes) gained, it is not possible to assess whether expenditure in a particular area is efficient.

The “cost savings” of either fully or partially preventing a given disease are, to a large extent, illusory. Assuming all the costs attributable to a given disease could be measured accurately and that adequate prevention were introduced, the cost savings from using cost of illness calculations are likely to be overestimated. Firstly, few diseases can be eradicated, so the total costs of treatment will not be saved. Secondly, when prevention fails certain capital investments, such as clinics, will continue to be required to treat those patients who still have the disease, so the marginal cost savings will be less than the average suggested by cost of illness studies. Finally, although treatment costs may be high, the costs of prevention could easily be much greater and a cost of illness study gives no information on prevention costs.
A further argument against the use of cost of illness studies as an aid to prioritising resources is that a high cost condition is not necessarily amenable to treatment by current medical technology. In contrast, a condition which presents a low cost to society may be fully amenable to low cost prevention, leading to high individual health gains. For example, because its incidence is low, untreated phenylketonuria, which leads to severe learning disability, will not present a great financial burden to society compared with, say, breast cancer. However, prevention is simple and inexpensive and the health gain to the individual is great. Thus, cost of illness studies may divert decision makers’ attention away from areas where important health gains can be made at low cost.

Thus, although widely undertaken, cost of illness studies add little to the creation of an efficient healthcare system. Current research efforts into costs of illness would be better focused on undertaking economic evaluations, such as a cost effectiveness analyses, which involve assessing both costs and outcomes.6

References
Costing in economic evaluation

James Raftery, director.
Health Economics Facility, University of Birmingham, Birmingham B15 2RT

As argued in previous notes, the perspective of an economic evaluationsocietal or confined to some organisationaffects the range of both outcomes and costs that should be included. The case has been made for taking a societal perspective, particularly for publicly funded programmes, but the range of resource effects that are potentially relevant is enormous. As with a stone dropped in a large pool of water the ripples in theory go on indefinitely, and some have spill-over effects. For practical purposes, however, one would want to measure only those occurring within a certain radius. The perspective adopted affects the "resource frame" of the study. The societal perspective implies taking a very large frame, while that of the public sector, or a particular organisation, implies a much reduced frame.

Three stages can be usefully distinguished in costing: identification, measurement, and valuation. Identification consists of listing the likely resource effects of the intervention as comprehensively as possible so that decisions can be made about the frame of the studythat is, which effects might reasonably be excluded. This decision depends on the perspective of the study. For example, a pharmaceutical company might restrict its perspective to the interests of the company. Even within a societal perspective, the decision about what to include in costing can be influenced by whether or not one wants to provide results which are comparable with other studies; for costing quality adjusted life years (QALYs) standardised approaches for "reference cases" have been proposed.2 Within a trial, restricting analysis to comparing the two or more arms may enable resource headings (non-healthcare charges, patients' and carers' time) common to each to be omitted. While this yields useful results, it restricts the comparability of the results to that service.

Measurement refers to the measuring of the resource changes included in the study. Typically these will be amounts of labour input or outputs (bed days, time in theatre, prescriptions) but may also include patients' or carers' time.

The final stage refers to valuation of these resource effects. If prices exist, and can be assumed to reflect costs, then these can be multiplied by the relevant units of service use to yield total costs, such as x bed days multiplied by £y per day. However, two problems complicate this simple picture: prices often do not exist for the relevant changes, and available prices may not reflect the societal value of resources. Economic theory suggests that prices will reflect resource values only under conditions of perfectly competitive marketsa situation which applies in part in the economy generally and hardly at all in health care. Economists have developed a range of methods for adjusting prices so that they better reflect resource use in studies with a societal perspective.

Two strategies can be usefully distinguished in measuring and valuation: microcosting and gross costing. Microcosting refers to detailed analysis of the changes in resource use due to a particular intervention, similar to the time and motion studies. Such detailed, bottom up, collection of data on resource use may be necessary when changes are being made to existing services (adding an extra stage or test). With microcosting, valuation use is also likely to require customised work as prices are unlikely to be available. Although many analysts favour microcosting, it tends to be costly and runs the risk of being specific to particular contexts.
Gross or top down costing allocates a total budget to specific services such as hospital stays or doctors' visits according to rules. The simplicity of top down costing may be offset by a lack of sensitivity, which in turn depends on the type of routine data available. The choice between microcosting and gross costing depends on the needs of the analysis. Many studies use a mixture of the two, using microcosting for the direct costs of the intervention, and gross costing for other costs. Costs which are incurred long after the intervention (such as admissions in subsequent years) will, when discounted, be greatly reduced in value, which suggests gross costing for these.

The data sources available for costing depend on whether microcosting or gross costing has been adopted. Microcosting often relies on wage rates to value staff time. The use of national pay scales in the NHS provides readily available data. Drug prices are published in the British National Formulary; hospital inpatient costs at the level of healthcare resource groups are published annually but with limitations; and the Department of Health funds an annual publication of the unit costs in community care.

Finally, the costs and benefits of different approaches to costing have to be assessed. Though detailed and comprehensive microcosting is desirable in principle, practicality restricts it. For that reason alone, economists should arguably be involved in the early stages of study design.

References