

Cost-effectiveness analysis and efficient use of the pharmaceutical budget: the key role of clinical pharmacologists

Richard Edlin, Jeff Round, Claire Hulme & Christopher McCabe

Academic Unit of Health Economics, University of Leeds, Leeds, UK

Correspondence

Professor Christopher McCabe, Academic Unit of Health Economics, Leeds Institute of Health Sciences, Charles Thackrah Building, University of Leeds, Leeds LS2 9LJ, UK.

Tel.: +44 113 343 6989

Fax: +44 113 343 3470

E-mail: c.mccabe@leeds.ac.uk

Keywords

budget, cost-effectiveness, efficiency, opportunity costs

Received

12 August 2009

Accepted

26 September 2009

The purpose of this paper is to provide information about cost-effectiveness analysis and the roles of clinical pharmacologists generally in providing efficient health care. The paper highlights the potential consequences of 'off-label prescribing' and 'indication creep' behaviour given slower growth (or potential cuts) in the NHS budget. This paper highlights the key roles of clinical pharmacologists in delivering an efficient health care system when resources are allocated using cost-effectiveness analyses. It describes what cost-effectiveness analysis (CEA) is and how incremental cost-effectiveness ratios (ICERs) are used to identify efficient options. After outlining the theoretical framework within which using CEA can promote the efficient allocation of the health care budget, it considers the place of disinvestment within achieving efficient resource allocation. Clinical pharmacologists are argued to be critical to providing improved population health under CEA-based resource allocation processes because of their roles in implementation and disinvestment. Given that the challenges facing the United Kingdom National Health Service (NHS) are likely to increase, this paper sets out the stark choices facing clinical pharmacologists.

Introduction

Economic evaluations refer to formal comparisons of alternative actions in terms of their costs and benefits [1]. Cost-effectiveness analysis (CEA) describes those analyses in which costs are measured in monetary terms with outcomes measured using a common unit of effect. The choice of outcome measure in a specific analysis should be determined by the decision which it is intended to inform. When the decision is concerned with alternative ways of achieving the same clinical effect such as reducing blood pressure, a clinical measure of outcome is appropriate, e.g. reduction in mmHg. A more general outcome is required when the alternatives are expected to provide multiple effects, or where different kinds of treatment are compared. This may be a clinical measure such as life years; however, CEA increasingly uses the Quality Adjusted Life Year (QALY) to measure effectiveness. CEA is also known as Cost Utility Analysis where it uses QALYs as its unit of effect.

The concept of the QALY is quite simple. Health care can improve the quality of an individual's life, their life expectancy or both. Therefore, a measure of health that captures changes across both domains of effect can be used to compare any two or more health care

interventions. The QALY weights life years lived by the quality of life experienced in each time period.

The terms health technologies and interventions refer to all possible types of health care and are used interchangeably, in the literature and in this paper. Measuring the impact of a health technology on life expectancy is relatively straightforward. Measuring its impact on quality of life is less so. It requires a framework for describing health related quality of life and a method for attaching weights to each state in the descriptive system. When cost-effectiveness analyses use QALYs to measure impact of a health technology, quality of life is measured using preference weights. These weights express the relative desirability of a health state on a scale anchored at 1 and 0; where 1 is the value attached variously to 'full health', 'perfect health' or 'best imaginable health' and 0 is the value attached to health states considered equivalent to being dead. Various methods are available for obtaining these weights, but discussion of these is outside the remit of this paper. They are well reviewed elsewhere [2, 3] and we note only that those methods using choices to determine preferences are considered superior to those that do not.

It is important to note that cost-effectiveness analysis considers all the costs associated with the delivery of a

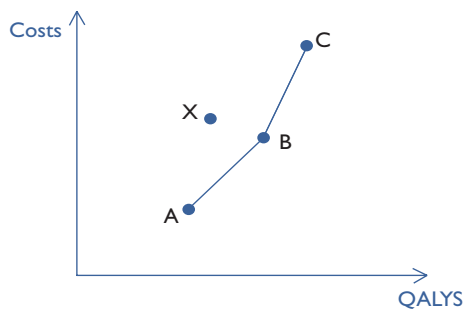


Figure 1
Incremental cost-effectiveness ratios

health technology to patients, not merely acquisition cost (price). It can also consider the costs (and savings) that accrue to patients and their carers, and even the impact upon economic activity. The methods for doing this are well described elsewhere [1].

Once the costs and effects of the alternative health technologies have been measured, they are compared by calculating the incremental cost-effectiveness ratio (ICER). The ICER is calculated as the difference in the expected costs of the two health technologies, divided by the difference in their expected effects (QALYs). If an existing health technology costs C_{OLD} and produces Q_{OLD} QALYs, and a new health technology B costs C_{NEW} and produces Q_{NEW} QALYs then the ICER is:

$$ICER = \frac{C_{NEW} - C_{OLD}}{Q_{NEW} - Q_{OLD}}$$

where more than two health technologies are compared, the ICER compares each option with its next best alternative in outcome terms. For instance, if we consider health technologies A, B and C in Figure 1, we first compare A with B and then B with C. Here, the slope of the lines between these points are the ICER figures and we expect these lines to have higher slopes as the number of QALYs gained increases. Where a health technology (e.g. X) has higher costs and lower benefits than a feasible health technology (such as B) it is said to be 'dominated' and will never be the best option available.

The cost-effectiveness threshold is the maximum value for an ICER that the decision maker will accept as the basis for a positive reimbursement decision. As we do not know the precise value of an ICER, we can only estimate the threshold. As such, it is important to consider the nature and magnitude of the uncertainty around each ICER. This is in turn determined by the uncertainty in each of the individual components of costs and outcomes. There is an extensive literature on appropriate methods for incorporating uncertainty into the analysis, and this is again detailed elsewhere [4].

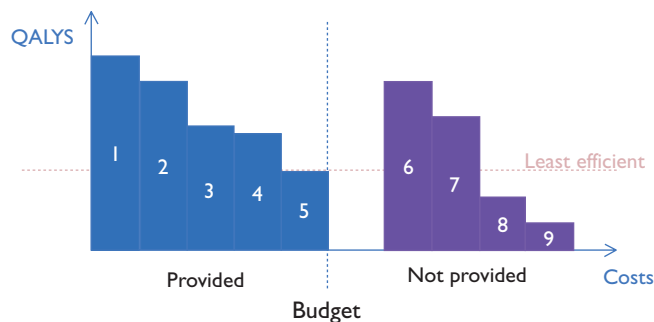


Figure 2
Cost-effectiveness across currently funded (1–5) and unfunded (6–9) health technologies

Uncertainty can be presented to the decision maker using a number of techniques including scatterplots on the cost-effectiveness plane, cost-effectiveness ellipses, cost-effectiveness acceptability curves (CEAC) and frontiers (CEAF). A CEAC plots the probability that the health technology is cost-effective over a range of possible values for the cost-effectiveness threshold [5]. The CEAF identifies the cost-effective health technology over a range of threshold values and is particularly useful when a comparison involves more than two health technologies [6]. The cost-effectiveness threshold is discussed in more detail in the next section.

Cost-effectiveness analyses can be conducted as an integral part of the randomized controlled trials that provide the evidence for the safety and effectiveness of a technology. However, normally such trials are multinational and whilst the effectiveness evidence for the new technology is likely to be transferable, comparator technologies, resource use and even quality of life impact can be health system specific. As a result, evidence from the randomized controlled trials is often combined with 'local data' in a cost-effectiveness model, to produce health system-specific cost-effectiveness estimates.

How can CEA be used to allocate a limited budget?

The framework within which ICERs are used to inform resource allocation decisions assumes a fixed health care budget spent to maximize the output from that budget [7, 8].

Figure 2 illustrates the choice facing the decision maker. There is a portfolio of health technologies (1–5) provided by the health system and the budget is completely spent on these interventions. Each health technology produces a certain amount of incremental health per £1000 spent upon them (measured on the vertical axis). The health technologies are ordered, from left to right, accord-

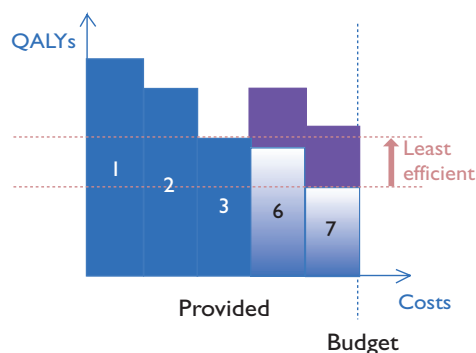


Figure 3

Efficiency gains following decisions to disinvest (4–5) and invest (6–7) within a fixed budget

ing to how much incremental health they produce per £1000 of expenditure, i.e. their efficiency. The cost-effectiveness threshold is determined by the cost-effectiveness of the least efficient health technology in the current NHS portfolio, including its formulary. (The lower the cost-effectiveness threshold, the more health was produced by the least efficient intervention (5).)

On the right hand side of Figure 2, there are a number of health technologies that the health system does not provide (6–9). In deciding whether to incorporate one of these health technologies into the health system's portfolio, the decision maker considers whether it would produce more health per £1000 spent than the least efficient health technology currently provided. If it does, it should displace the least efficient health technology in the current portfolio.

Two of the new health technologies (8, 9) provide less health per £1000 spent than the least efficient current health technology. Making these available would reduce average population health. The other two technologies (6, 7) produce more health per £1000 spent than the least efficient current health technology. Making these available in the NHS will increase population health as long as they are funded at the expense of less efficient NHS health technologies.

The shaded area above 6 and 7 in Figure 3 illustrates the population health gain associated with incorporating these more efficient technologies. Here, health technologies 4 and 5 have been replaced by interventions 6 and 7. More efficient health technologies can now be funded because less efficient health technologies have been identified and withdrawn. Note also that as the health system becomes more efficient, any new health technology must be more efficient before it justifies inclusion in the NHS portfolio. Whereas before, health technologies only had to be more efficient than health technology 5 to qualify, now they have to be more efficient than health technology 3.

In this way, a health system charged with allocating a budget that it cannot change has a threshold that reflects

the health technologies available to the system, and not a general willingness to pay across society. The threshold is an estimate of the efficiency of the least efficient thing that the system does. Of course, where a budget is increased the threshold rises since a higher budget allows the system to provide less efficient interventions that are currently unaffordable.

Within the UK, NICE has used the same basic threshold figure for 10 years, which suggests a rough equivalency between the effects of greater efficiencies and the effects of a higher budget to the NHS. The results of empirical analyses of the threshold have been broadly consistent with the NICE value. Whilst we could expect the NHS to continue to make efficiency savings, it seems highly unlikely that the NHS budget will continue to grow at recent rates. As such, the threshold might be expected to fall in the medium to long-term. This effect is compounded by moves to place greater values on health gains at the end of life and on innovative health technologies, since these tend to be both expensive and not generally cost-effective, using current criteria. Spending money on these things means that the budget is exhausted more quickly, so that the true threshold in the NHS is currently falling.

This means that even some technologies that are currently considered cost-effective may not be in future, and it makes it much more likely that any new interventions must be directly funded by disinvesting from less efficient actions. However, when the NHS provides tens of thousands of health technologies, how do we know which things are the least efficient, and how can we establish just how efficient they are? Currently NICE decisions assume that the threshold is around £20 000 per QALY, although they will approve health technologies with ICERs above this range as additional factors such as certainty and equity considerations are considered. This threshold figure is as much informed by judgement as by science. NICE have been encouraged to be careful in their use of these additional considerations as the lack of information about interventions that are displaced to fund the new treatments, means they cannot assess whether the patients who bear the opportunity cost have fewer, the same or more 'special' characteristics than the beneficiaries of the positive decision [8].

Disinvestment and efficient resource allocation

It is simply not feasible to identify the cost-effectiveness of all health technologies that the NHS provides. An alternative approach, which estimates the marginal health gain associated with a marginal increase in the health care budget, has been demonstrated using programme budget data. Martin *et al.* adopted this approach using NHS Programme budget data [9]. Their results indicate that the threshold varies significantly across different health care

specialties. They estimate that the threshold ICER in oncology is in the region of £28 000 per QALY but only around £8000 per QALY in respiratory disease. One implication of these results is that the overall efficiency of new health technologies depends in part on what must be displaced to fund it: population health improvements are likely when disinvesting from oncology, but unlikely when disinvesting in respiratory care. Indeed, it is probable that any technologies funded by reducing respiratory care have actually worsened overall population health.

An alternative approach to matching investment and disinvestment with a fixed budget is Programme Budgeting and Marginal Analysis (PBMA). PBMA has been used sporadically since the mid-1970s, and its principles are well established [10]. PBMA characterizes the decision problem using three questions:

- 1 If the current budget is fixed, could we increase the total benefit from its expenditure by redeploying resources from one programme to another?
- 2 If the budget were to increase, how should we allocate the additional resources to get the best value from them?
- 3 If the budget is to be cut, where should the reduction take place so as to minimize the benefit loss?

The first step in PBMA is to group the health system's activities into programmes. Programmes frequently reflect clinical specialties, such as oncology, cardiology and musculoskeletal medicine. The budget and activities of each programme are then identified. Particular attention is paid to finding those activities that produce the least value within a programme. Each programme is also invited to identify the most valuable additional interventions that could be implemented with additional resources.

Experience with PBMA suggests that it is resource intensive and rarely, if ever, adopted as an organisation's long-term framework for making resource allocation decisions. However, the value of its focus on identifying the marginal activities for disinvestment and investment is acknowledged. If new investments are funded through disinvesting from those interventions that it is easy to disinvest from, rather than those that are least valuable, the health of the population can be harmed rather than promoted. Making the right disinvestment decision is as important as making the right investment decision.

The key role of clinical pharmacologists

Cost-effectiveness analyses, and by extension the decisions that rely on them, typically assume that interventions are used in line with best practice and only in the patients for whom they are indicated. Where recommended treatments are not used, are used but not according to best

practice, or are given to wider groups than the treatments will not be as cost-effective as expected. Inefficient prescribing, for example, will mean that the expected benefits for a new pharmaceutical will not be realized. This can seriously complicate decisions, and particularly where they rely on evidence from RCTs in which prescribing behaviour is tightly controlled.

Thus, NICE, and similar organisations are completely dependent upon the front line health care professionals adhering to the detail of their recommendations, in order to improve population health and allow confidence in future recommendations. It is questionable whether clinical pharmacologists always recognize their centrality in this regard. Whilst NICE has an overt commitment to clinical freedom over-riding guidance [11], this freedom is an acknowledgement that individual cases differ and not *carte blanche* to extend guidance inappropriately. Guidance on how clinical pharmacologists, and indeed all clinical staff, should balance their responsibility to the individual patient and patients, plural, is limited [12]. Greater engagement from Royal Colleges and the General Medical Council with regards to this difficult issue seems warranted.

If 'indication creep' and 'off label' prescribing [13] continue then it may become difficult to avoid highly bureaucratic and resource intensive systems that identify precisely who receives which treatment. Such systems allow micro level billing and auditing and are typical within privately funded systems. They ensure that health technologies are only provided to those patients for whom they are approved, but entail significant costs and so divert resources away from front line care. Experience from commercially financed healthcare systems suggests that such systems can be a significant cost driver. The business case for implementing such systems is dependent upon the degree of divergence between best practice (which underlies the commissioning decisions) and actual practice. There is an unavoidable tension between a clinician's commitment to do the very best for the individual patient and their responsibility to protect and promote the health of all patients. Whilst liberal interpretation of guidance may be in the interests of individual patients, it is not in the interests of the health of the population.

As NHS budgets become tighter over the coming months and years, resolving the tension between the interests of the individual and the population may well determine whether the NHS is sustainable. It will also have a huge impact on the way we allocate health care resources. If clinical pharmacologists prioritize the individual over the population then we are likely to see an increasingly adversarial and managerial model of health care delivery, with the associated costs decreasing the resources available to help patients.

Even if bureaucratic systems are not introduced, a simpler and more brutal alternative exists. If decision makers cannot rely on clinical pharmacologists to pre-

scribe as recommended, then judgements must be made taking inappropriate prescribing into account. In such cases, decisions makers would need to establish that new and expensive health technologies are superior almost regardless of the decisions made by clinical pharmacologists. In this case, it is difficult to see new therapies becoming routinely available to anyone within the NHS. The limits placed on clinical pharmacologists and the ability of the NHS to aid patients would again be drastically curtailed.

Discussion and conclusions

Cost-effectiveness analysis is a prominent and increasingly influential tool for decision makers to guide policy. A form of economic evaluation, cost-effectiveness analysis compares outcomes across a series of health technologies to ask what the most appropriate use of resources should be if the aim of the health care system is to maximize population health. Decisions under CEA are relatively simple: those new health technologies providing health more efficiently than current practice should be adopted in favour of existing health technologies that provide health less efficiently.

CEA is not the only way of making decisions, with other frameworks for decision making possibly allowing consideration of additional factors, including equity. However, these frameworks involve additional analytical tools and typically require more and different information than CEA. The purpose of this paper is not to consider the merits of alternative frameworks but rather to consider the specific issues of efficiency. Regardless of the specific framework used to aid decision makers, efficiency will require the identification of more efficient and less efficient uses of resources, and suggests that the former should be preferred to the latter. The stark conclusions of this paper are not specific to CEA.

Increasingly stringent budget constraints, increased NHS efficiency, and policies requiring NICE to recommend selected, more expensive health technologies pose a serious challenge within the NHS. This challenge will mean that some health technologies currently considered to be efficient will be stopped. NICE will soon have to place more stringent limits on recommending interventions not considered to be special cases. Only greater efficiencies within the NHS, and particularly those resulting from the clinical behaviour of professionals, can increase the ability of NICE to recommend new health technologies by freeing up additional resources.

Further efficiency savings are likely to require greater reliance on re-imburement recommendations. Where clinical pharmacologists act instead against re-imburement recommendations in a single patient's interest, this benefits the patient in the short term only and harms other patients as a result. This is particularly the case for expensive health technologies in which eligibility criteria

are typically tight, since a single inappropriate prescribing decision impacts upon a large number of people who all stand to lose potential access to other health technologies.

Given this, it is probably inevitable that clinical pharmacologists will face more stringent limits on their behaviour. The choice is a stark one: either clinical pharmacologists decide en masse to follow recommendations, or they may be forced to do so by a system that necessarily involves far greater restrictions on their activities. If this choice is made voluntarily, it can be done without decreasing the budget available for health care and can allow for rational decision making. However, if professionals must be coerced, the cost of this coercion is likely to involve reduced budgets for health care or draconian re-imburement systems that set extremely stringent conditions for new therapies.

Competing interests

C. McCabe has worked extensively with all stakeholders in the NHS Resource allocation process. Thus any change in practice will impact positively on some of his collaborators and negatively on others. There are no other competing interests to declare.

No specific source of funding was used in the preparation of this manuscript and no ethical approval was necessary.

REFERENCES

- 1 Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. *Methods for the economic evaluation of health care programmes*, 3rd edn. Oxford: Oxford University Press, 2005.
- 2 Brazier JE, Ratcliffe J, Tsuchiya A, Salomon J. *Measuring and valuing health benefits for economic evaluation*. Oxford: Oxford University Press, 2006.
- 3 Arnold D, Girling A, Stevens A, Lilford R. Comparison of direct and indirect methods of estimating health state utilities for resource allocation: review and empirical analysis. *Br Med J* 2009; 339: b2688.
- 4 Briggs AH, Claxton K, Sculpher M. *Decision modelling for health economic evaluation*. Oxford: Oxford University Press, 2006.
- 5 Fenwick E, Byford S. A guide to cost-effectiveness acceptability curves. *Br J Psychiatry* 2005; 187: 106–8.
- 6 Fenwick E, Claxton K, Sculpher M. Representing uncertainty: the role of cost-effectiveness acceptability curves. *Health Econ* 2001; 10: 779–87.
- 7 Culyer AJ, McCabe C, Briggs AH, Claxton K, Buxton M, Akehurst R, Sculpher M, Brazier J. Searching for a threshold, not setting one: the role of the National Institute for Health and Clinical Excellence. *J Health Serv Res Policy* 2007; 12: 56–28.

- 8** McCabe C, Claxton K, Culyer AJ. The NICE cost-effectiveness threshold: what it is and what that means. *Pharmacoeconomics* 2008; 26: 733–44.
- 9** Martin S, Rice N, Smith PC. The link between health care spending and health outcomes for the new English Primary Care Trusts, CHE Research Paper 42; 2008. Centre for Health Economics, University of York.
- 10** Mitton C, Donaldson C. Twenty-five years of programme budgeting and marginal analysis in the health sector, 1974–1999. *J Health Serv Res Policy* 2001; 6: 239–48.
- 11** National Institute of Health and Clinical Excellence. Legal context of NICE guidance. Available at http://www.nice.org.uk/media/8BD/2B/Legal_context_nice_guidance.pdf (last accessed 1 August 2009).
- 12** General Medical Council. *Guidance for Doctors: Management for Doctors*. London: GMC, 2006.
- 13** Hébert PC, Stanbrook M. Indication creep: physician beware. *Can Med Assoc J* 2007; 177: 697.