

Crossing the fourth hurdle

Michael D. Rawlins

National Institute for Health and Clinical Excellence, 71 High Holborn, London WC1V 6NA, UK

Correspondence

Professor Michael D. Rawlins, Chairman,
National Institute for health and Clinical
Excellence, 71 High Holborn, London
WC1V 6NA, UK.

Tel.: +44 084 5003 7780

Fax: +44 084 5003 7784

E-mail: michael.rawlins@nice.org.uk

Keywords

comparative effectiveness, cost
effectiveness, direct comparisons,
distributive justice, indirect comparisons,
priority setting

Received

5 January 2012

Accepted

29 February 2012

Accepted Article Published Online

7 March 2012

The fourth hurdle, the requirement that pharmaceutical manufacturers can demonstrate that their new products represent good value for money as well as being of good quality, effective and safe, is increasingly being required by healthcare systems. In crossing this 'fourth' hurdle, companies will usually need to demonstrate that their products are more effective than relevant comparators and that the increased cost is offset by the enhanced benefits. Decision makers, however, must draw their conclusions not only on the basis of the underpinning science but also on the social values of the people they serve.

Introduction

In order to obtain approval for new medicines to be placed on the market, national drug regulatory authorities require pharmaceutical companies to show that these products are:

- of appropriate pharmaceutical quality
- effective in the indications for which they are to be used and
- that they are safe in relation to their efficacy.

In recent years these three 'hurdles' have been supplemented, in an increasing number of healthcare systems, by a requirement for manufacturers to show that their new products are also cost effective. This requirement, although not part of the regulatory process, has been described, colloquially, as the 'fourth' hurdle. Countries requiring economic evaluations of new products, before recommending their use include Australia, Canada, Finland, France, Germany, New Zealand and Sweden as well as England, Wales and Scotland [1].

The need for healthcare systems to undertake economic assessments of new pharmaceutical products (as well as other interventions) derives from a simple fact. The resources (combining private and public expenditures)

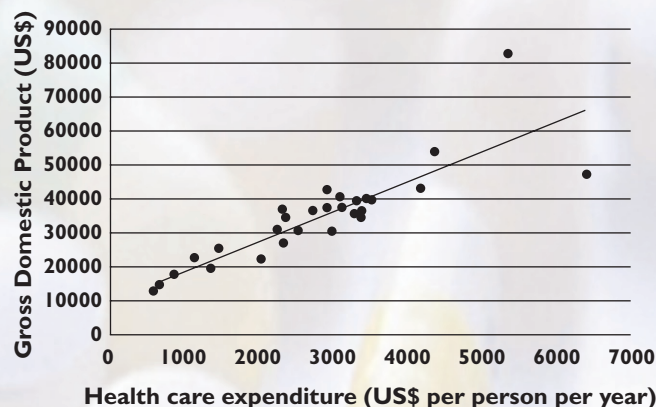


Figure 1

Relationship between OECD member states' expenditure on health care and their Gross Domestic Products (both expressed as US\$ at purchasing power parity)

that developed countries devote to healthcare are closely related to their wealth as reflected by their Gross Domestic Products (Figure 1). It is obvious from Figure 1 that interventions that are cost effective in wealthier countries will not necessarily be cost effective in poorer ones. In reality

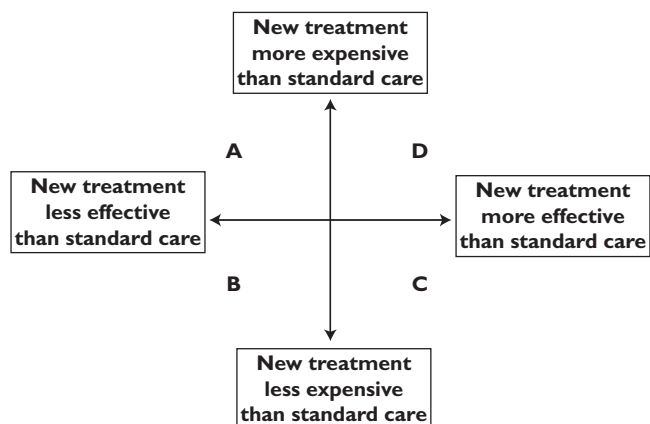


Figure 2

Cost effectiveness plane depicting the relationship between comparative clinical effectiveness and comparative cost effectiveness

no country, not even the richest, is able to provide its citizens with all their healthcare needs and priorities have to be made. Examining the cost effectiveness of new pharmaceuticals, especially in countries where healthcare is largely or exclusively publicly funded, is just one part of a wider attempt to control healthcare expenditure.

The evaluation of a product's cost effectiveness is often depicted as a 'cost effectiveness plane' (Figure 2). Products in quadrant A (more expensive but less effective) would obviously be unattractive for any healthcare system. Products in quadrant B (less expensive but less effective) would pose problems for decision makers but are very unusual with new pharmaceutical products. Products in quadrant C (less expensive but more effective) would obviously be highly desirable interventions but, again, are uncommon with new products. Most new pharmaceuticals fall in quadrant D (more effective but also more expensive).

Comparative clinical effectiveness

Scrutiny of Figure 2 shows that, for the purposes of an economic evaluation, the effectiveness of a new pharmaceutical needs to be compared with one (or more) relevant comparators [2]. In doing so, two questions have to be answered. First, what are the relevant comparators (sometimes called 'competing interventions')? Second, how much more effective is the new product when compared with the relevant comparator(s)?

Defining the relevant comparator(s)

The definition of the relevant comparator(s) is often difficult and may vary both between and within countries. Lack of agreement among clinical experts, practising in similar clinical environments, is surprisingly common even within

a single healthcare system such as Britain's National Health Service (NHS). Difficulties also arise when the relevant comparator(s) are products that are not authorized (licensed) in the circumstances for which the new treatment will be used (i.e. 'off label'). Even more difficult issues are raised if the relevant comparator(s) is not authorized, for any indication, in the form or formulation to be used as a comparator (i.e. 'unlicensed').

The use of 'off label' and 'unlicensed' comparators is especially fraught in assessing the use of medicines in children. Many widely used pharmaceutical products are rightly, and widely, used in children in the absence of licensed indications or formulations. To do otherwise would deprive children of appropriate care. Despite these inherent problems, decisions have to be made after consultation with relevant experts, patient organizations and the relevant pharmaceutical manufacturer.

Evidence synthesis

Direct comparisons The most obvious approach to comparing the effects of two (or more) alternative treatments can be obtained from the results of head-to-head trials. Where the results of two or more independent studies are available, appropriate meta-analyses will provide the best estimate of comparative effectiveness.

Indirect comparisons In practice, in part because of the preference of drug regulatory authorities for placebo controlled trials rather than active comparator controlled trials for licensing purposes, direct comparisons of new pharmaceutical products are often unavailable at the time they are launched.

Indirect comparisons involve estimating the comparative effectiveness of two or more alternative treatments in the absence of head-to-head trials. Thus, if treatments A and B have both been studied in separate placebo controlled trials, it is possible to impute the effects of A vs. B. In doing so, it is important to avoid the so-called 'naïve' approach involving the pooling of data across treatment arms [3]. Rather, the comparisons should be 'adjusted' by using trials that have an intervention in common such as placebo. Thus, rather than crudely comparing the results of treatment A vs. treatment B (the 'naïve' approach), the results of the differences between treatment A and placebo should be compared with the differences between treatment B and placebo (the 'adjusted' approach). Adjusted indirect comparisons potentially overcome the problem of differential prognoses between study participants in the trials, as well as preserving the benefits of randomization.

Mixed treatment comparisons Mixed treatment comparisons (also known as network meta-analyses or multiple treatment meta-analyses) involve the synthesis of both direct (if available) and all indirect comparisons within a single model [4]. When conducted appropriately, these can

provide estimates of each pair-wise treatment effect without breaking randomization. The approach is particularly useful when there are few or no direct comparisons.

The use of indirect and multiple treatment comparisons are relatively recent methodological innovations and their adoption was initially controversial. Multiple treatment comparisons, moreover, usually require the construction of complex statistical models requiring considerable expertise in their construction and interpretation if the results are to be reliable and informative [5–8]. They are now used extensively by organizations, such as the National Institute for Health and Clinical Excellence (NICE) in the appraisal of both new and established interventions [9]. They have also been successfully used, for example, in assessing the comparative effectiveness of established treatments for depression [10] and acute mania [11].

Generalizability

The results of randomized controlled trials provide indications as to how a product performs in ideal circumstances and assess the efficacy of the product. In these circumstances the investigators are usually experts in the field, the patients are usually a relatively homogeneous population and the duration of treatment is often relatively brief when compared with that likely in the real world. The extent, therefore, to which these results will be relevant to use in normal clinical practice (i.e. ‘effectiveness’) often remains uncertain [2].

Apart from so-called ‘pragmatic trials’, which attempt to include a wider cohort of patients, there is no easy solution and it is, to date, largely a matter of judgement for decision-makers. This is especially difficult in assessing the clinical effectiveness of treatments likely to be used for very much longer periods than could possibly be encompassed in a randomized controlled trial. It is in this area, perhaps above all others, that clinical pharmacologists can contribute most by virtue of their inherent knowledge of factors such as the development of tolerance, the appearance of long latency adverse reactions and the natural history of the underlying condition.

Comparative cost effectiveness

Economic evaluations in healthcare require clear definition of the underlying principles in addition to the technical elements involved in estimating cost effectiveness [2].

Underpinning economic principles

At the start of any economic evaluation of healthcare interventions economists are likely to be constrained by three underpinning principles. These often reflect ethical, legal, social or political issues that have been pre-determined by governments, legislatures or others.

Perspective The perspective, or viewpoint, from which the economic analysis will be undertaken, may be confined

solely to the costs and benefits to the healthcare system. This is the perspective, as mandated in its Statutory Instruments, used by NICE. The perspective could, however, be wider and encompass the costs and benefits that fall on public expenditure or, as in Sweden, the costs and benefits falling on society as a whole. The reason why NICE’s remit is narrow, covering only the costs and benefits to the health and the personal social services, is because the Institute is legally part of the NHS and cannot have powers beyond those devolved by parliament to the British NHS.

Economic assessment The nature of the economic assessment used by health economists could be based purely on the budgetary impact and affordability of the intervention, but is most frequently (as is the case with NICE) undertaken using cost utility analysis. In this the incremental costs (and savings), together with the incremental benefits (expressed as quality-adjusted life years (QALY) gained), are used to calculate the incremental cost effectiveness ratio (the ICER). This is described in greater detail below.

Distributive justice This is the term, used by political philosophers, when discussing what is appropriate, or right, for the allocation of goods in a society [12]. Utilitarians seek to maximize the health of the population as a whole. Egalitarians, in so far as is possible, want each individual to receive a fair share of the available opportunities.

Both approaches have their merits. Utilitarianism emphasizes the importance of efficiency and egalitarianism reminds us of the importance of fairness. Each, though, poses difficulties that conflict with the moral convictions of many [13]. Utilitarianism, for example, does little to remedy health inequalities; while egalitarianism, in attempting to distinguish between what is fair and unfair, as well as between what is unfair and unfortunate, lacks clarity. NICE, like many other comparable bodies, attempts to achieve a balance on a case-by-case basis between these sometimes conflicting approaches to distributive justice.

Cost effectiveness

Cost utility analysis is the preferred approach, by NICE and many other similar organizations, to cost effectiveness analysis. This technique allows the cost effectiveness of one intervention, for one condition, to be compared with the costs and benefits of another intervention for a different condition [2].

Costs and savings In estimating the costs accruing to the use of a particular pharmaceutical product, account is obviously taken of its acquisition costs, in comparison with the relevant comparator(s). The use of a pharmaceutical product may also consume other healthcare resources. These might include the cost of administration and special monitoring requirements as well as the costs of treating adverse effects. Depending on the economic perspective, as discussed above, account might also be given to indirect

costs such as time off work, the payment of sickness benefits and for temporary help at a patient's home. Savings, too, compared with alternative treatments will also be taken into account.

Benefits In cost utility analysis the benefits, the improved quality of life achieved by the use of an intervention, are assessed from the increase in health 'utility' [2]. This scores the quality of life from zero (dead) to one (perfect health). If, on this scale, an intervention results in a change in the utility score of (say) 0.4 to 0.7, the gain in utility will be 0.3. The gain is then multiplied by the years for treatment is 'enjoyed' to yield the QALY gained. The incremental QALY gained is the difference between the QALY gained by one intervention and the QALY gained by the alternative one [2].

Incremental cost effectiveness ratio (ICER) The ICER is the incremental net costs of using a new treatment (A) vs. the costs of the relevant alternative treatment (B), divided by the incremental QALY gains of the two same treatments:

$$ICER = \frac{Cost_A - Cost_B}{QALY\ gained_A - QALY\ gained_B}$$

On some occasions treatment A is less costly but more effective. Rather than calculating a negative ICER economists merely state that treatment A is 'dominant' [2]. In most instances, however, the new treatment A is more costly but more effective. Decision makers must therefore determine whether the magnitude of the ICER represents good (or at least reasonable) value for money for the healthcare system.

Cost effectiveness thresholds

The acceptability of a particular product's ICER, for its use in a particular indication, is determined by the 'threshold' distinguishing cost effective, from cost ineffective, interventions. NICE rejects the use of an absolute (fixed) threshold for four reasons [14, 15]:

- There is a weak empirical basis for deciding at what value a threshold should be set.
- A fixed threshold would imply that efficiency has an absolute priority over other objectives such as fairness.
- Pharmaceutical companies are monopolies and a fixed threshold would discourage price competition.
- Rigid adherence to a fixed threshold would imply full acceptance of the calculations underpinning the estimates of cost effectiveness. It would therefore remove discretion in assessing costs and benefits when modelling has reached its limits.

Rather than apply an arbitrary fixed threshold, NICE makes its decisions on a case-by-case basis as shown stylistically in Figure 3. As the ICER increases so does the

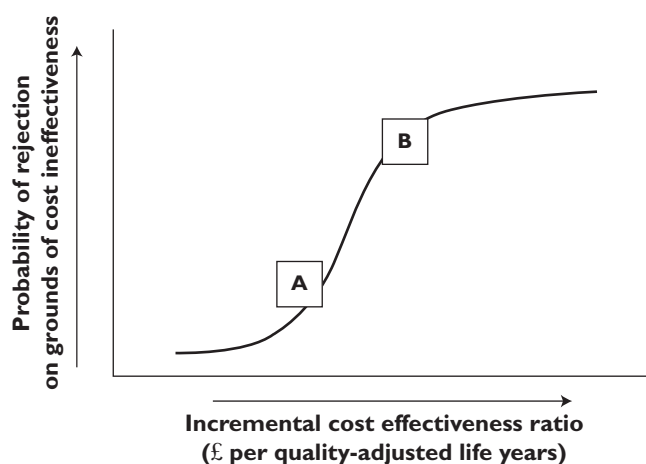


Figure 3

Relation between the likelihood of a technology being considered cost ineffective vs. the incremental cost effectiveness ratio (ICER)

probability of rejection on grounds of cost ineffectiveness. The controversial issue is at what values the ICERs should be at inflections A and B in Figure 3.

In the collective judgement of NICE's health economic advisers, interventions with an ICER less than £20 000 per QALY gained (inflection A in Figure 3) should usually be considered as cost effective. Special reasons would be needed to accept technologies with ICERs above £30 000 per QALY gained (inflection B in Figure 3). The use of products with ICERs above £30 000 per QALY gained would, if adopted, be likely to deny other patients, with other conditions, cost effective care.

Decision making

In drawing their conclusions about the cost effectiveness of a new pharmaceutical product, for use in any publicly funded healthcare system, decision makers must make both scientific and social value judgements.

Scientific judgements

The scientific evidence supporting the use of a new product is always incomplete and decision makers have to make judgements about the available data [2]. The evidence base is always deficient in some way. Decision makers will need to make judgements about the reliability of the evidence base and consider the impact, on any conclusions, of the inevitable imperfections and uncertainties. Do the intermediate outcomes really reflect the likely ultimate outcome? Is it reasonable to assume that the effectiveness of products, assessed in relatively short term studies, will be sustained over many years? When an intervention is cost ineffective overall, can the available evi-

dence justify conclusions about its use in apparently cost effective sub-groups? Can the results of the premarketing randomized controlled trials be generalized to the heterogeneous population of patients who will be treated in routine clinical practice? And, very importantly, has the quality of life been adequately captured with the tools used in the relevant studies? If not, what is the likely effect on the ICER?

These scientific judgements are best made by the members of NICE's advisory bodies who are appointed, in large part, because of their expertise at scrutinizing and interpreting the technical and clinical data.

Social value judgements

Social value judgements relate to society rather than basic or clinical science. They take account of the ethical principles, preferences, culture and aspirations that should underpin the nature and extent of the care provided by healthcare systems [14]. In balancing the sometimes conflicting tenets of utilitarianism and egalitarianism NICE and its advisory bodies must take into account social values. Such values include:

- Should the NHS be prepared to pay more, to extent a child's life by a year, compared with older people?
- Do treatments that prolong life, at the end of life, have a special importance that warrants increased expenditure?
- Should 'more generous' consideration be given to treatments for serious conditions?
- Should a healthcare system be prepared to pay 'premium prices' to treat very rare conditions?

However, neither the Institute's board, nor the members of its advisory bodies, have any special legitimacy to impose their own social values on the NHS. Rather, these should reflect the values of the public who own, and ultimately fund, the NHS [12]. The Institute's social values have largely been developed by its Citizens Council, formed from members of the public, whose advice has been collated into guidelines for the use of its advisory bodies [16].

NICE's advisory bodies attempt to be explicit about the social values that have been adopted in particular forms of NICE guidance. At the present these are expressed qualitatively and there is no attempt to weight, quantitatively, specific elements. As methodologies develop it may be possible, in the future, to incorporate social values, quantitatively, using multi-criteria decision analysis. Even then, because of the difficulties of identifying all possible and relevant future social values, qualitative elements are always likely to exist.

Conclusions

Crossing the fourth hurdle is part evaluation, assessment and appraisal of the relevant clinical science. It should also,

though, attempt to align decisions about the availability of (especially new) pharmaceuticals, in a healthcare system, with the social values of the paying public. The ultimate decisions are culturally and context specific and they may not readily cross trans-national boundaries. They may sometimes be controversial and disappoint special interest groups in the pharmaceutical industry, in the healthcare professions and amongst specific patient-advocacy organizations. Nevertheless such stakeholders have a right to be heard, to present their arguments and to 'have their say'. They cannot always, though, have their way. To do so would deprive other groups of patients, many of whom lack influential advocates, access to cost effective care.

Crossing the fourth hurdle has been seen by some as a disincentive to investment in the development of innovative new pharmaceutical products. There may be some truth in this but it is an inescapable fact that, especially in the current economic climate, it is impossible for any country, anywhere, to provide medicines for all its peoples irrespective of their costs. In reality, NICE has only rejected around 12% treatment condition cost comparisons it has made. A new product cannot, in my view, be considered to be innovative if it is unaffordable to the health care system for which it has been developed.

The fourth hurdle may also, at least potentially, have adverse effects on so-called 'incremental innovation'. This term describes modest improvements in a product, or product class, over existing treatments. Incremental innovation can indeed offer additional benefits to patients; but in such instances it is appropriate for health care systems to pay a premium price only in proportion to the marginal benefits that the product brings.

Like it or not the fourth hurdle is here to stay and has been introduced, albeit in different ways, by most developed countries with health care systems based on the principles of social solidarity. The pharmaceutical industry realizes this and is trying to adapt its business models accordingly. It still requires drug regulatory authorities, though, to understand that increasing the regulatory burdens on the life sciences industries generally endangers the future of these industries and the prospects for patients around the world. More of that, perhaps, for another day!

Competing Interests

There are no competing interests to declare.

REFERENCES

- 1 International Society for Pharmacoeconomics and Outcomes Research. Pharmacoeconomic Guidelines around the World. Available at <http://www.ispor.org/PEguidelines/index.asp> (last accessed 29 March 2012).

- 2** Rawlins MD. *Therapeutics, Evidence and Decision-Making*. London: Hodder Arnold, 2011.
- 3** Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R, Bradburn M, Eastwood AJ. Indirect comparisons of competing interventions. *Health Technol Assess* 2005; 9.
- 4** Caldwell DM, Ades AE, Higgins JPT. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005; 331: 897–900.
- 5** Song F, Loke YK, Walsh T, Glenny A-M, Eastwood AJ, Altman DG. Methodological problems in the use of indirect comparisons in evaluating healthcare interventions: survey of published systematic reviews. *BMJ* 2009; 338: b1147; doi:10.1136/bmj.b1147.
- 6** Donegan S, Williamson P, Gamble C, Tudur-Smith C. Indirect comparisons: a review of reporting and methodological quality. *PLoS ONE* 2010; 5: e11054.
- 7** Ades AE, Mavranzouli I, Dias S, Welton NJ, Whittington C, Kendall T. Network meta-analysis in competing risk outcomes. *Value Health* 2010; 13: 976–83.
- 8** Song F, Xiong T, Parekh-Bhurke S, Loke YK, Sutton AJ, Eastwood AJ, Holland R, Chen Y-F, Glenny A-M, Deeks JJ, Altman DG. Inconsistency between direct and indirect comparisons of competing interventions: meta-epidemiological study. *BMJ* 2011; 343: d4909; doi: 10.1136/bmj.d4909.
- 9** National Institute for Health and Clinical Excellence. *Guide to the Methods of Technology Appraisal*. London: National Institute for Health and Clinical Excellence, 2008.
- 10** Cipriani A, Furukawa TA, Salanti G, Geddes JR, Higgins JPT, Churchill R, Watanabe N, Nakagawa A, Omori IM, McGuire H, Tansella M, Barbui C. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *Lancet* 2011; 373: 746–58.
- 11** Cipriani A, Barbui C, Salanti G, Rendell J, Brown R, Stockton S, Purgata M, Spinelli LM, Goodwin GM, Geddes JR. Comparative efficacy and acceptability of antimanic drugs in acute mania: a multiple treatments meta-analysis. *Lancet* 2011; 378: 1306–15.
- 12** Rawlins MD. Pharmacopolitics and deliberative democracy. *Clin Med* 2005; 5: 471–5.
- 13** Beauchamp TL, Childress JF. *Principles of Biomedical Ethics*. Oxford and New York: Oxford University Press, 2001.
- 14** Rawlins MD, Culyer AJ. National Institute for Clinical Excellence and its value judgements. *BMJ* 2004; 329: 224–7.
- 15** Rawlins M, Barnett D, Stevens A. Pharmacoeconomics: NICE's approach to decision-making. *Br J Clin Pharmacol* 2010; 70: 346–9.
- 16** National Institute for Health and Clinical Excellence. *Social Value Judgements: Principles for the Development of Nice's Guidance*. London: National Institute for Health and Clinical Excellence, 2008.