NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Briefing paper for the Methods Working Party on the Cost Effectiveness Threshold

The briefing paper is written by members of the Institute's Decision Support Unit. It is intended to provide a brief summary of the issues that are proposed for discussion by the Working Party to inform an update to the Institute's Guide to Methods of Technology Appraisal. It is not intended to reflect a comprehensive or systematic review of the literature. The views presented in this paper are those of the authors and do not reflect the views of the Institute.

The briefing paper will be circulated to the members of the Method's Review Working Party, the group responsible for updating the guide.

For further details regarding the update of the Guide to the Methods of Technology Appraisal please visit the NICE website at www.nice.org.uk/TAMethodsReview.

1 Review of the 'Guide to Methods of Technology Appraisal'

The Institute is reviewing the 'Guide to the methods of technology appraisal', which underpins the technology appraisal programme.

The original Methods Guide was published in February 2001, and a revised version was published in April 2004. The Methods Guide provides an overview of the principles and methods used by the Institute in assessing health technologies. It is a guide for all organisations considering the submission of evidence to the technology appraisal programme and describes appraisal methodology.

The current 'Guide to methods of technology appraisal' is available from the NICE website at http://www.nice.org.uk/page.aspx?o=201974.

The review of the Methods Guide will take place between April and October 2007. As part of the process, a number of workshops will be held to help identify those parts of the Guide that require updating. These workshops will involve a range of stakeholders, including methods experts, patient representatives, industry representatives, NHS staff and NICE technology appraisal committee members.

The revised draft of the Methods Guide will be available for a 3-month public consultation, expected to begin in November 2007. We encourage all interested parties to take part in this consultation.

2 Background

2.1 Relevance of the Topic to NICE

The Institute is charged with considering both the effectiveness and cost effectiveness of treatments and then with making recommendations as to their provision within the National Health Service (NHS). Cost Effectiveness Analysis (CEA) assesses two or more alternative courses of action in terms of their costs and benefits. The comparison is summarised using the expected Incremental Cost Effectiveness Ratio (ICER). This is a measure of the additional cost per additional unit of health gain produced by one intervention compared to another. The Institute's preferred form of cost effectiveness analysis uses the Quality Adjusted Life Year (QALY) to describe the outcome of each intervention. By extension, the Institute's preferred form of ICER is the Cost per QALY gained.

The cost effectiveness threshold is often referred to as society's willingness to pay for an additional unit of health gain (QALY); i.e. if the ICER for a specific intervention is less than the cost effectiveness threshold, then society, or its agents, will be willing to fund it, and conversely if the ICER is greater than the threshold society will not be willing to fund it. If NICE uses a cost effectiveness threshold that is too high, then its guidance will promote inefficient uses NHS resources. If however, the Institute's threshold is too low, then valuable interventions will not be adopted by the NHS and the NHS will not make the best use of available resources.

This conception of the cost effectiveness threshold wrongly presumes, however, that there is a direct link between positive recommendations by the NICE Technology Appraisal Programme and the NHS budget. NICE is charged with making recommendations on how the NHS should improve the health of the population it serves with available resources. It is explicitly not allowed to make recommendations on whether the government should make additional resources available to the NHS to fund a specific intervention. [1] Although the 3 month Department of Health (DoH) mandate implies that the DoH considers the expected cost of implementing NICE Appraisal Guidance in setting the NHS budget.

The objective of this paper is to review the current state of knowledge regarding the cost effectiveness threshold, the principles of its use in health care resource allocation decisions and any arguments for and against changing the threshold from the current £20,000.

Section 3 summarises the current statements in the methods guide regarding the value and of use of the cost effectiveness threshold. Section 4 critically reviews the relevant literature on the cost effectiveness threshold. Section 5 draws key implications of using an ICER threshold to promote population health gain from the NHS budget. Section 6 considers the empirical evidence

on the value of the threshold for the NHS and what searching activities the Institute might consider to inform its understanding and utilisation of the cost effectiveness threshold. Section 7 attempts to summarise the key issues around the cost effectiveness threshold for the Methods Working Party to consider.

3 What the current methods guide says

The current methods guide contains a number of statements that are relevant to the consideration of the cost effectiveness threshold. In chapter 6 it states:

"The Appraisal Committee does not use a fixed ICER threshold above which a technology would automatically be defined as not cost effective or below which it would. Given the fixed budget of the NHS, the appropriate threshold is that of the opportunity cost of programmes displaced by new, more costly technologies. However, estimating this threshold would require complete information about the costs and QALYs from all competing healthcare programmes and the Committee does not have this information. Furthermore, the threshold will change over time as the budget for healthcare changes. Although the use of a threshold is inappropriate, comparisons of the most plausible ICER of a particular technology compared with other programmes that are currently funded are possible and are a legitimate reference for the Committee."[15]

This statement acknowledges the importance of the considering the opportunity cost of implementing new treatments given a fixed threshold – whilst conversely suggesting that, since the data required to estimate the threshold quantitatively are not available, it is inappropriate to use a threshold.

The Guide then goes on to consider the use of cost effectiveness analysis in practice in terms of a range of possible factors to take into account at or around an empirically plausible threshold:.

"Below a most plausible ICER of £20,000/QALY, judgements about the acceptability of a technology as an effective use of NHS resources are based primarily on the cost-effectiveness estimate. Above a most plausible ICER of £20,000/QALY, judgements about the acceptability of the technology as an effective use of NHS resources are more likely to make more explicit reference to factors including:

- the degree of uncertainty surrounding the calculation of ICERs
- the innovative nature of the technology
- the particular features of the condition and population receiving the technology
- where appropriate, the wider societal costs and benefits."

This approach echoes ideas advanced by Akehurst in 2002.[2] The statement suggests that there is a "plausible" cost effectiveness threshold of £20,000 per QALY. When cost effectiveness ratios for a treatment exceed this threshold the Appraisal Committee should consider (a) whether the characteristics of the condition or population receiving the treatment would lead them to value the health gain produced by the intervention more highly than the estimate made in the analysis; (b) whether innovative characteristics of the intervention are such that the appropriate weighing of the Secretary of State's instruction to consider innovation would lead to positive recommendation, despite the excess opportunity cost from a pure efficiency perspective; (c) whether other benefits to society, outside of those considered by the cost effectiveness analysis, are such that it is 'socially desirable' for the treatment to be made available; and (d) whether the uncertainty in the cost effectiveness estimate is such that the risk of a false negative decision suggests the wisdom

of treating the estimate as not significantly different from the threshold value (a consideration first outlined by Akehurst)[2]

It is worth noting that the utilisation of these additional criteria are not inconsistent with the operation of an explicit cost effectiveness threshold and also with both the instructions of the Secretary of State to the Institute and the research literature on society's preferences regarding the allocation of scarce health care resources. However, as discussed below, the use of a cost effectiveness threshold does have implications for how these considerations should be put into practice.

4 The cost effectiveness threshold

Three approaches to establishing a cost effectiveness threshold have been proposed:[3]

1. Inferring thresholds from previous decisions;
2. Setting thresholds in order to determine the optimal budget; and
3. Setting the threshold to exhaust and optimal budget.

4.1 Inferring the threshold from previous decisions

Rawlins and Culyer published the first attempt to infer NICE's cost effectiveness threshold from published decisions. 1 [4] On the basis of an essentially qualitative analysis of published decisions they observed that 'NICE rejects the use of an absolute threshold…. As the incremental cost effectiveness ratio increases, the likelihood of rejection increases.'

They illustrate their argument with an S-shaped curve (Reproduced as Figure 1) and state that 'NICE and its advisory bodies have taken the view that

inflexion A occurs at around £5000 - £15,000/QALY and inflexion B at around £25,000 - £35,000/QALY.'

This description of a threshold range echoes Akehurst's concept of a 'smudge'[2] where the smudge 'reflects increasing discomfort as the cost goes higher and higher'.

Around the same time, Devlin and Parkin published a statistical estimation of NICE's threshold, utilizing data from 51 observable yes/no decisions. The authors used logistic regression analysis to identify the most likely value of the ICER threshold; concluding that it is 'somewhat higher than the £20,000 - £30,000 which NICE has publicly identified.'[5]

There are two disadvantages to utilising previous decisions to determine the cost effectiveness threshold. The first is that is not necessarily desirable for current decisions to be constrained to the same decision criteria as previously used. As explained below, (see Section 5) consistency of decision rule (the cost effectiveness threshold) may be in conflict with consistency of objective (maximising health gain).

The second is a practical issue. Thus, as Devlin and Parkin acknowlege, the consistency approach assumes that all factors relevant to the decisions and all evidence regarding each factor,was fully captured in the dataset. [5] If this condition is not met, then the ICER derived from the analysis is unlikely to represent the decision rule utilised by decision makers and thus its adoption would lead to inconsistency rather than consistency with previous decision making.

4.2 Setting the Optimal Budget

Setting a threshold to determine the optimal budget entails specifying a value for health gain (e.g. incremental cost per QALY) and then providing all interventions for which the ICER is at or below this value. The health care budget would then be the total value of the resources required to implement this policy. An important characteristic of this approach is that setting the threshold effectively determines the budget, for a given portfolio of interventions.[6] It also commits the state to increase the budget so long as the ICERs for new interventions fall below the threshold value.

The key question raised by this approach is, 'How should the value of a QALY be determined?'. Three approaches can be identified in the literature.

- Establishing society's willingness to pay for health gain.
- Setting the threshold equal to Gross Domestic Product per Head of Population.
- Setting the threshold equal to the value of life/health used in other public sector decisions.

4.2.1 Willingness to pay for health gain

NICE, in conjunction with the NHS R&D Programme, has funded two research projects in collaboration with the NHS R&D Methodology Programme. They are designed to examine what value society attaches to health gain and how it varies according to the characteristics of the recipient.

The link between the general public's stated willingness to pay for health gain, and the threshold for use by NICE ought to be is unclear The budget that Parliament approves for the NHS can be interpreted as implying a value of health gain. The process of Parliamentary approval may be considered to

provide such a value with a social legitimacy that the estimated willingness to pay lacks.

4.2.2 Set the threshold equal to per capita Gross Domestic Product

Williams suggested that a 'common sense' value for the threshold would be to set it equal to the per capita GDP.[7] The intuitive appeal of this proposal is that if every member of society were to be given a fair share of nation's wealth, they would receive the per capita GDP. The maximum they could therefore spend on health gain in any one year would therefore be the per Capita GDP. If expenditure exceeds this value either the nation is spending more than it earns or some people are receiving less than their fair share. However, there are a number of problems with this approach. Firstly, it assumes that a nation would wish to devote all of its wealth to health care. Secondly, its ignores the fact that it is possible for the average cost effectiveness to be at or below this threshold when the cost effectiveness of marginal programmes is markedly higher. The cost per QALY of the marginal programme is likely to be considerably higher than the average due to diminishing marginal returns. Thus, the proposal does not recognise the nature of the decision facing NICE; i.e. whether or not is efficient to stop the current least efficient intervention in order to fund the new intervention.

4.2.3 Set the threshold equal to the value of life/health utilised in other public sector decision processes.

Loomes suggests that the threshold should be set at a level which is consistent with the value attached to a life in other parts of the public sector.[8] He observed that the average value of a year of life in full health implied by the Department of Transport's valuations for saving a life was in the region of £30,000. Whilst health and life are the primary (although not sole) objectives of the National Health Service, they are not the primary objectives of other public sector activities. A socially legitimate set of relative values of these

different objectives is implicit in the budgets voted to each activity by Parliament. An estimated value based upon a series of reference activities chosen by NICE, would lack such legitimacy.

Loomes describes a number of significant methodological challenges to producing robust monetary values of health gain. Outside of the transport and health fields, research on the value of lives saved or health gained as the result of specific policies is not collected. In the absence of such data it is difficult to see how to set the threshold by reference to values in other areas of public policy, and as Culyer observes 'NICE simply does not have (and nor is it mandated to acquire) the kind of information about outputs in non-health sectors that it would need to form necessary judgements about the marginal costs and benefits of health spending versus spending in other areas of public services'.[3]

4.3 Setting thresholds to exhaust a budget optimally

From the beginning, NICE's use of cost effectiveness analysis has been perceived as a means of promoting the efficient use of available NHS resources. Commenting upon the establishment of the NICE Appraisal Programme, Maynard and Hutton state that 'Budget holders can use cost effectiveness analysis or cost utility analysis to maximize health gain from allocation of limited resources.'[9] Devlin also describes the cost effectiveness threshold as a means of ensuring that available NHS funds are spent on those treatments that produce the greatest health.[10] The ICER threshold ought thus to be the cost per QALY of the least efficient funded treatment (i.e. the intervention with the highest cost per QALY). For a new intervention to be provided it must be more efficient than the least efficient currently funded intervention and ought to displace it in whole or part in the portfolio of funded interventions. Devlin argues that the challenge of identifying the

marginal interventions is too great and therefore an alternative rationale for determining the cost effectiveness threshold is required. It is difficult to understand why Devlin, having recognised what the threshold represents, goes on to argue that because it will be difficult to identify the correct value for the threshold, we must pretend that it is something else.[10]

Birch and Gafni have repeatedly argued that decision makers cannot maximize health gain from limited resources by using ICERs in isolation from information on budget impact. To do so, they say, is a recipe for 'continued expansion of expenditure. [11] [12] [13] [14] Their argument is that the opportunity cost of a positive decision is determined by the total budgetary impact, not the incremental cost effectiveness ratio. It is possible for the total budget for an existing health intervention to be less than that for the new intervention even though the new intervention has a lower ICER. If the new intervention is mandated on the basis of the ICER alone, then 'extra funds would need to be found'.[12] [A]

Birch and Gafni's analysis implicitly assumes that there is no relationship between the ICER threshold and the budget, and also that the budget is fixed. However, as we shall see below, the ICER threshold is dynamic. The greater the facility of the appraisal programme to take account of this fact, the lower the risk of Birch and Gafni's fears becoming reality.

In 2007 Culyer and colleagues described a model of NICE as a threshold searcher. [6] The model starts from the same point as Devlin where the interventions provided by the NHS are ranked in descending order according to the marginal health gain.[10] The cost effectiveness threshold is the ICER for the least efficient intervention currently provided.

[A] Birch and Gafni also argue that for the use of CEA/CUA to be consistent with welfare economics assumptions of constant returns to scale and complete divisibility must hold. Whilst nothing is completely divisible, the pragmatic importance of divisibility in the scale of resources required for particular technologies relative to total NHS budget is not clear.

Figures 2a and b reproduced from Culyer et al, illustrate the situation where there are some interventions not provided by the NHS that are more efficient than some interventions that are provided.[6] The function of NICE, is postulated to be to substitute more efficient interventions for less efficient ones. It does this through specifying a 'working' cost effectiveness threshold, reflecting the Institute's estimate of the ICER of the least cost effective activity undertaken by the NHS. This working estimate is drawn from (a) the incomplete evidence base on the cost effectiveness of interventions that the NHS does provide; and (b) stakeholders knowledge of the likely value of funded interventions for which formal evaluations are not available. [6 15] Over time this 'working' ICER should be updated reflecting developments in the published evidence base and evidence on the efficiency of the interventions that are disinvested from, in order to fund the recommended interventions.

It is worth noting that the opportunity cost of NICE recommendations (given an unchanging NHS budget) is the value of the health gain lost from activities that are actually displaced. There may be a divergence between the activities that should have been displaced on pure efficiency grounds and the activities that are displaced in practice. Where the activities that are likely to be displaced are not the least efficient, the estimated threshold should be identified on the basis of what is likely to be displaced, otherwise the opportunity cost of positive recommendations is likely to exceed the value of the health gain they produce.

The local commissioning structure of the NHS means that the success of the NICE Technology Appraisal Programme in promoting population health gain depends on the quality of local commissioning. If local commissioners consistently disinvest from relatively efficient activities in order to fund NICE recommendations the net effect will be a reduction in population health. Effective communication and co-operation between local commissioners and

the Institute around investment and disinvestment activities is therefore essential if both parties are to meet their obligations.

Culyer and colleagues argue that NICE needs both to monitor commissioners' disinvestment activity and to consider making disinvestment as a kind of threshold "searcher".[6] Empirical work on actual value of the health gain displaced is now starting to appear, Smith and colleagues (ref); and NICE themselves have commissioned exploratory work from Parkin and colleagues at City University on displacement. (Personal communication, Kalipso Chalkidou, July 2007).

4.3.1 A comment on Birch and Gafni using the 'Threshold Searcher' model

The model described by Culyer et al is useful for exploring Birch and Gafni's critique of the use of ICERs to make resource allocation decisions.[11]

The cost effectiveness threshold is the inverse of the marginal health gain per unit of expenditure of the least efficient intervention in current use. In Figure 2, we can see that the substitution of a more for a less efficient intervention causes the marginal health gain of the least efficient intervention to rise. The cost effectiveness threshold has therefore decreased. The next candidate intervention will need to be even more efficient in order to justify its inclusion as a funded intervention. This is the case even if the budget impact of the substitution is neutral – i.e. the total cost of the new intervention is identical to the total cost of the displaced intervention. The cost effectiveness threshold is thus better seen as dynamic, rather than as a fixed parameter.

If the total cost of the new intervention is greater than the marginal intervention, then positive recommendations will require a greater degree of disinvestment , until the budgetary impact of the substitution is neutral and

the budget constraint holds. This means that the cost effectiveness threshold for an intervention with a large budgetary impact should be lower than that for an intervention with a small impact. In this way, cost effectiveness allocation processes based on a threshold can fully capture the opportunity cost of both positive and negative investment recommendations.

Culyer et al's model also implies that the threshold should increase if the budget is increased,– i.e. less efficient interventions can be incorporated into the portfolio of treatments provided by the NHS (assuming that the initial threshold was correctly specified for the initial budget.). In times of rapid expansion of the NHS budget, such as we have seen over the past 7 years, the countervailing effects of the implementation of new treatments and increases in the budget may have made the adoption of a fixed cost effectiveness threshold less inappropriate. However, when the budget is fixed or growth is less than the net budget impact of investment and disinvestment decisions, the cost effectiveness threshold should fall to reflect the increased efficiency of the marginal intervention.

Using this model it is possible to see that the cost effectiveness threshold can capture the opportunity cost of positive recommendations and therefore that the NICE resource allocation processes are not, contrary to B&G, necessarily a recipe for an ever increasing NHS budget.[11] .

4.4 Summary

The budget of the National Health Service is set by Parliament. NICE is charged 'to appraise the clinical benefits and costs of such health care interventions as may be notified by the secretary of state or the National Assembly for Wales….and to reach a judgement as to whether, on balance, this intervention can be recommended as a cost effective use of NHS and PSS resources.' It is clear that NICE is not mandated to determine the budget of

the NHS. Since setting a threshold independent of the budget is logically equivalent to determining the budget, NICE is not mandated to do so[6] Thus the appropriate model for considering NICE's cost effectiveness threshold is that of identifying the threshold that is consistent with promoting the efficient use of a fixed budget.

5 Implications of using the cost effectiveness threshold to exhaust a fixed budget optimally

If we accept that the appropriate use of the cost effectiveness threshold by the Institute is to optimally exhaust a fixed budget, what, if any, are the implications for the methods of the NICE Appraisal Process and the AC's approach to decision making?

5.1 The Threshold ICER and Innovation

Figure 3 [16] shows the total health gain to the NHS population under 3 scenarios. Consider an intervention costing £20,000 per patient. At this price the ICER is below the cost effectiveness threshold and the net health benefit of the intervention is 1 QALY per person. If the price were £40,000, the ICER is exactly equal to the threshold and at this point the net benefit from the new intervention is 0; the loss of health from displaced technologies being the same as the gain.

However, if the new treatment is more effective than existing treatments, then setting price at a level that produces an ICER equal to the cost effectiveness threshold means that the full value of the innovation (greater efficacy) is received by the manufacturer. As the manufacturers are profit maximisers they will wish to price as close to this point as possible. Note, if the price

exceeds £40,000, then the net benefit becomes negative; i.e. total population health would be reduced by the provision of the new intervention.

While it may be important for manufacturers to appropriate a share of the value of the innovation in order to give an incentive for investment in research and development, concern for dynamic efficiency does not mean that they ought to take it all. Firstly, the public sector subsidises research and development in a number of ways, through publicly funded research, tax incentives and research infrastructure investment. Therefore, even if society was unconcerned about who benefits from innovation (NHS patients or the pharmaceutical industry) it would not be efficient to allow full appropriation of the value of innovation by the manufacturer. In other markets where innovation is protected society simply permits monopoly rents during patent protection but does not allow full appropriation by, for example, facilitating perfect price discrimination (see below). Finally it is appropriate for society to be concerned about who receives the benefits and it is reasonable that at least some of the benefits of innovation should accrue to NHS patients.

However, such concerns as this are somewhat tangential to those of the Institute's. Of more direct relevance is the use of innovation as an argument for recommending interventions having ICERs above the threshold. When the ICER is close to, or at the threshold value, the full value of the innovation is already being paid to the manufacturer. To recommend an intervention even though the ICER is above the threshold is to pay more for the innovation than it is worth. Promoting population health is consistent with recommending treatments with ICERs substantially below the threshold. Taking account of the value of innovation in health care, as per the directions of the Secretary of State, is consistent with approving interventions upto the cost effectiveness threshold.

5.2 Monitoring and recommending disinvestment within a threshold searcher model.

The current methods guide does not explicitly recommend a cost effectiveness threshold on the grounds that the correct figure could not be known. However, the model discussed above suggests that it may be legitimate for the Institute to utilise an explicit threshold within the context of a search activity.[6] A crucial part of this search activity is to identify activities for disinvestment or, when growth monies, are being considered, to identify the formerly planned investments that are not to be undertaken in order to fund NICE recommendations.

There has been little research on disinvestment in the NHS. NICE has commissioned research from Brunel and City Universities; [17] however, this is not expected to report until September. If this research reports that the disinvestments tend to be more cost effective than the NICE recommended interventions, this will be a prima facie case for supposing that the current threshold is too high. However, it will not represent conclusive evidence, unless the decision processes used by commissioners embody cost effectiveness as a primary decision criterion.

A recent Faculty of Public Health Conference on commissioning saw several presentations suggesting that commissioners are adopting evidence based commissioning strategies incorporating cost effectiveness considerations.[18] The greater the degree of take-up of these processes, the more likely it seems that disinvestment will be mainly in the least efficient interventions.

Martin and colleagues examined the actual changes in budgets and health across Primary Care Trusts (PCTs) and estimated the average budget elasticity of health; this is the change in health that has resulted from marginal changes to programme budgets. Therefore it does provide an important empirical estimate of the cost effectiveness threshold. Within programme

budgets rather than across PCT budgets they found, for example, that the threshold for oncology is slightly in excess of £19,000; and that for cardiovascular disease is around £12,500. [19]

An alternative approach suggested by Culyer and colleagues is that NICE actively pursue disinvestment as well as investment recommendations.[6] Recently the Institute has started to explore the possibilities in this area.[20] In addition, external organisations have started to recommend that the NHS, via NICE or other routes, disinvests from activities that do not have a robust evidence base.[21]

Knowing the value of health gain foregone is essential if the appraisal programme is to be confident that it is contributing to increased population health. Knowing which interventions are actually displaced will also help NICE is to understand the distributional effects of its recommendations.

5.3 The dynamic threshold ICER

The Chair of NICE recently observed that the current threshold range has been utilised for 7 years and noted that the Methods Review Process would need to consider whether the range should change or remain the same. [22] To consider whether the threshold should change it is important to understand that (a) the threshold can be determined empirically and (b) the future changes in the cost effectiveness threshold are related to budget impact of current decisions.

5.3.1 Empirical estimates of the cost effectiveness threshold

Martin et al, have shown that it is possible to establish the value of the cost effectiveness threshold empirically.[19] This work focussed upon just two programme budgets; oncology and cardiovascular disease, however, these

also happen to be the two largest clinical areas in which NICE has issues appraisal guidance. As we have seen, the results make a prima facie case that the threshold should be set lower than £20,000 per QALY.

The method utilised by Martin et al is generalisible and there is a case for NICE to encourage its much wider application. NICE might also consider ways in which the results of such analyses could be used to update its cost effectiveness threshold regularly - possibly annually but certainly more frequently than has been the case to date.

5.3.2 Budget impact and the threshold ICER

To understand the relationship between the ICER threshold and budget impact, consider the situation in Table 1. Suppose that the health service has a total budget of £500,000. With this budget it currently purchases 5 interventions. The marginal health gain of each ranges from 12 to 4 QALYs. All treatments cost £1000 per person and each treatment is provided to 100 people. The least efficient treatment (treatment 5) has an ICER of £250 per QALY.

A sixth treatment is now brought to the market having a marginal health gain of 5.5 QALYs and an ICER of £182. Treatment 6 is substituted for Treatment 5, the least efficient of the current treatments. Should a Treatment 7 be launched it would need to have an ICER < £182 per QALY in order to represent a more efficient use of NHS resources than the current portfolio, (Treatments 1,2,3,4 and 6).

Table 1: The dynamic cost effectiveness threshold.

Intervention	Mhg (QALYs)	Cost	ICER	Prevalence	Total Cost	Total Budget
Treatment 1	12	1000	83	100	100000	500000
Treatment 2	10	1000	100	100	100000	
Treatment 3	8	1000	125	100	100000	
Treatment 4	6	1000	167	100	100000	
Treatment 5	4	1000	250	100	100000	
Treatment 6	5.5	1000	182	100	100000	

Now consider a scenario where the price of the Treatment 6 is not £1000, but £1200, producing an ICER of £218. Treatment 6 is still more cost effective than Treatment 5 (ICER=£250 per QALY) and thus, using an ICER decision rule, should be implemented. However, the total cost of implementing Treatment 6 is greater than the total cost of Treatment 5, under this scenario the relevant threshold (assuming indivisibility) is not £250 but £167; the ICER for Treatment 4; as funding Treatment 6 requires some disinvestment from Treatment 4 in addition to the disinvestment from Treatment 5. Alternatively, the budget could be increased by £20,000 to cover the increase in costs.

This second scenario illustrates the importance of considering whether the budgetary impact of the interventions that NICE recommends is marginal. When the budgetary impact is not marginal, then the relevant decision threshold will be lower. The Office of Fair Trading estimates that the NHS provides 25,000 various interventions with its budget of approximately £90 billion per annum.[23] The Treasury has estimated the total cost of NICE guidance to date to be in excess of £1 billion ; i.e. approximately twice the NHS deficit for 2005/6, which led to widespread moratoria on NHS activity. Whilst these figures do not prove that NICE's recommendations go beyond marginal expenditure, they certainly indicate caution in assuming that they are.

The Institute is charged with deciding priorities within existing resources. [1] Utilising an ICER-based decision rule, when resources are fixed requires that the ICER threshold is dynamic, the greater the budget impact of guidance to date, the greater the reduction in the ICER threshold required to ensure future decisions do not displace established interventions with less efficient ones.

5.4 Equity arguments with a threshold ICER

The current methods guide makes clear that when the ICER is above the threshold, the Appraisal Committee considers whether there are any characteristics of the patient group that would justify valuing the expected health gain more highly.[15] The only factor that is explicitly described is the 'degree of clinical need' – suggesting that greater weight will be given to health gain when the untreated health state is "severe".

The threshold represents the opportunity cost of the implementation i.e. the health gain foregone by other patients. While the threshold is critical to the determination of the most efficient i.e. health maximising use of NHS resources, the Appraisal Committee must also consider whether there is any ground in equity for weighting the health gains and losses of different people differentially or for recommending technologies with relatively high ICERS on grounds of their beneficial impact on equity..15

Return to the resource allocation scenario described above. Let the new treatment now cost £1600 and let the ICER be £291. The ICER suggests that the new treatment is not more efficient than Treatment 5, and thus should not be reimbursed. However, it turns out that the new treatment is for a severe condition and, for slightly under two thirds of patients failure to provide it will hasten the end of an already foreshortened life. The decision maker decides that health gain in these circumstances is worth 1.20 times the health gain in the reference case. The 'equity weighted ICER' is thus £242 per QALY

and Treatment 6 is substituted for Treatment 5 in the sub-group of patients for whom it is appropriate.

Table 2: Equity weighted ICERs

Intervention	Mhg (QALYs)	Cost	ICER	Prevalence	Total Cost	Total Budget
Treatment 1	12	1000	83	100	100000	500000
Treatment 2	10	1000	100	100	100000	
Treatment 3	8	1000	125	100	100000	
Treatment 4	6	1000	167	100	100000	
Treatment 5	4	1000	250	100	100000	
Treatment 6	5.5	1600	291	100	100000	
Equity Weight	**20%**					
Treatment 6	6.2	1600	242	62	99000	
Treatment 5	4.8	1000	208	100	100000	

Now consider the possibility that treatment 5 was for a different severe condition in the later stages of life. If this was the case, the equity weighted ICER for Treatment 5 would be 208 and it would no longer be efficient to reimburse Treatment 6 even using equity weighted ICERs.

Given the evidence that the majority of health care consumption takes place in the last year of life,[24] there is a substantial probability that patients who bear the opportunity cost of positive decisions will have characteristics which themselves *might* carry an equity weight – e.g. imminence of death, severity of condition, and lack of alternative treatment. While NICE cannot know ex ante which patients will bear the opportunity cost of its recommendations,[1 3 6 11 15] given that the majority of health care consumption takes place in the last year of life, the implicit current presumption that all patients carry an equal weight does not seem plausible.

Whilst NICE is unable to know which patients will bear the cost of its guidance, a formal disinvestment programme is one potential means of ensuring that displaced interventions are less efficient than the mandated

intervention and that the equity characteristics of those who benefit from guidance are weightier than those of the people bearing the opportunity cost.

5.5 Choice of health state values and the use of an ICER threshold.

The ICER threshold also has implications for the question of whose values should be used to value health gains. It is increasingly accepted that the use of General Population values may be inappropriate.[25] However if patient values are used then they need to be used to establish the value not only of the health gain foregone but also of the gain produced.

Suppose that patients value health gain 20% higher than the reference case assumption. Treatment 6 now becomes cost effective compared to Treatment 5. However, this is true only if it is assumed that the reference case values are an appropriate reflection of the views of those patients receiving Treatment 5, even whilst they are not for the patients who receive Treatment 6. It is not readily apparent why this should be so. If the patients who receive treatment 5 also value their health gain more highly (and to a similar degree) than the general public, then treatment 6 will not be an efficient use of resources compared to treatment 5.

Whilst there is an acceptance that general population values do not reflect the value of health gain to the recipients of that health gain, we do not have patients' values for all patients affected by guidance decisions. To use patient values to inform NICE recommendations is to value health gain to the patients who would gain from a positive decision differentially from the health of patients who will bear the opportunity cost of a positive decision.

6 Changing the Threshold ICER

When considering whether the threshold ICER should change, there is value in recognising that that the threshold is an empirical concept, rather than an issue of social value judgement; thus, where available, empirical evidence should be used to inform such considerations.

An alternative to changing the threshold value is to change the weight attached to the other factors considered alongside the estimates of the reference case ICER.

The empirical evidence from the study by Martin et al, indicates that the even the lower end of the current cost effectiveness range is too high and likely to lead to less efficient treatments being implemented at the cost of more efficient ones. Thus, there is a prima facie case for considering reducing the threshold. However, evidence on the threshold for a broader range of programmes would be desirable.

The insights provided by threshold searcher model indicate that:
(a) it is feasible to operate an explicit single threshold rather than the current range;
(b) innovation should be dropped as an argument for providing a positive recommendation to interventions when the ICER is greater than the threshold;
(c) the use of either equity weights or patient values to sustain a positive recommendation when the ICER is greater than the threshold ignores the equity characteristics and values of those patients who bear the opportunity cost and thereby risks being inequitable and inefficient; and
(d) that interventions that have a large budgetary impact should be required to have lower ICERs as they will displace more than the marginal activities.

The work of Claxton and colleagues suggests that it is wholly inappropriate to allow uncertainty in the evidence base to lead to a positive recommendation in when the reference case ICER is above the threshold. Indeed a full consideration of the consequences of an uncertain decision for the NHS suggests that positive guidance would require an ICER substantially below the threshold if the consequences of an uncertain decision are substantial. These consequences includes the potential sunk costs (e.g., investment in equipment, training and implementation) from guidance which is likely to change in the future and the impact of positive guidance on the prospects of acquiring certain types of evidence which may be needed to support NHS practice. However the important point is that the same threshold is needed to make these assessments.

7 Key Issues

The detailed consideration of the methodological and empirical literature relating to NICE' s cost effectiveness threshold has identified a number of key issues which the Working Party may wish to consider in deciding whether or not develop the current methods guidance statements regarding the use of a cost effectiveness threshold.

1. Should the Institute use a single threshold or a range?
2. Which criteria provide a legitimate basis for recommending a treatment with an ICER in excess of the threshold?
    a. Can innovation be dropped from the current list?
    b. Can uncertainty be dropped from the current list?
    c. What is the appropriate process for incorporating equity considerations

3. What, if any, empirical evidence should NICE seek to inform the specification of the threshold ICER?
    a. Programme budget and health outcome analyses
    b. Disinvestment monitoring
    c. Disinvestment guidance
4. How frequently should NICE reconsider the value of the threshold ICER?
5. Is the current empirical evidence sufficient to justify a reduction in the threshold ICER?
6. How, if at all, should the Appraisal Committee incorporate patient values into considerations around the threshold ICER?
7. When the budget impact of a positive recommendation is large (however defined) should the Appraisal Committee require an intervention to be more cost effective in order to receive a positive recommendation?

Figure 1: Relationship between cost effectiveness and probability of rejection (Rawlins and Culyer)

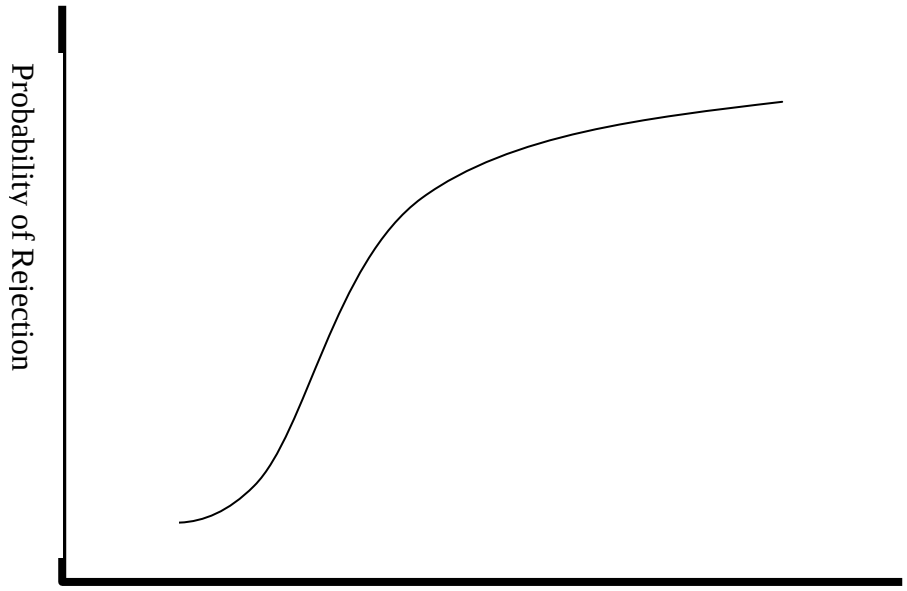


Figure 2: NICE as a Threshold Searcher

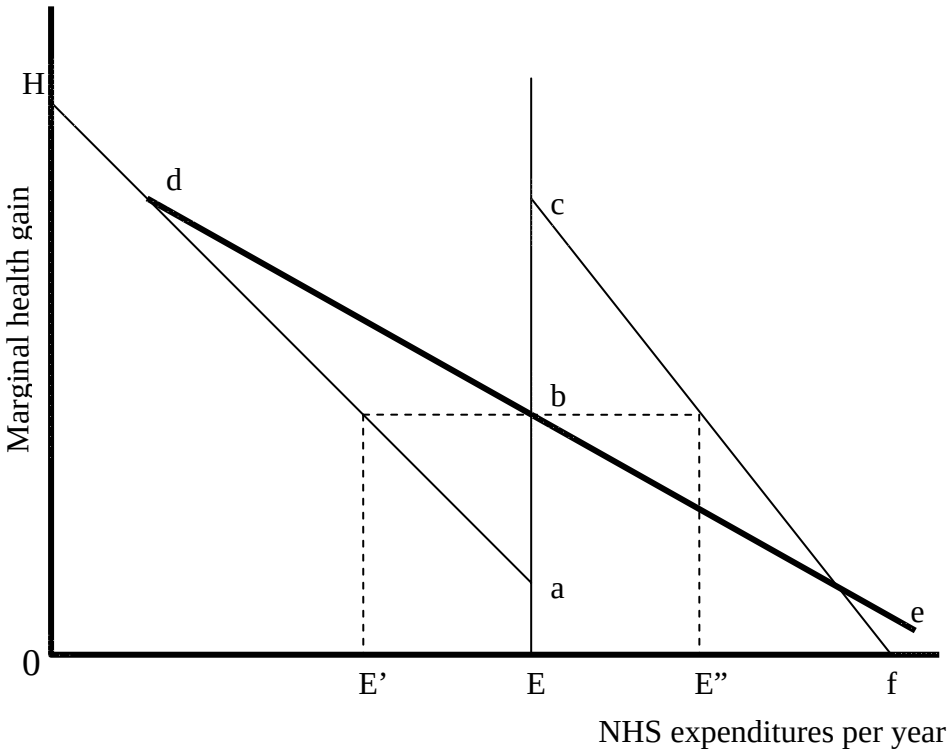
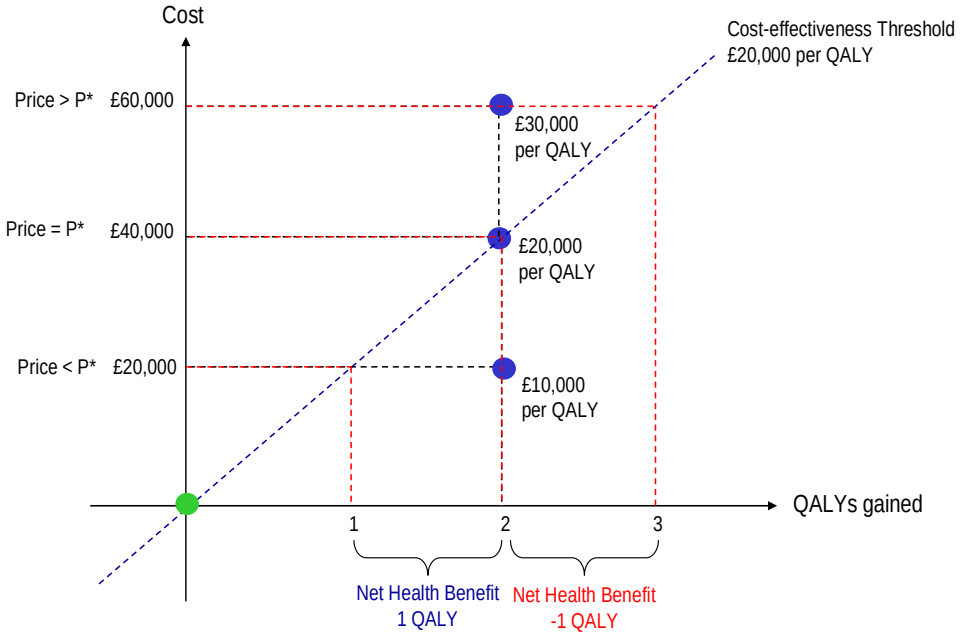


Figure 3: Threshold and Health Gain



References

[1] Rawlins MD. Culyer AJ National Institute for Clinical Excellence and its value judgements. BMJ 2004;329:224-227

[2] Akehurst R. in Towse, A., Pritchard, C. Devlin, N. Cost effectiveness thresholds: Economic and Ethical Issues. London Kings Fund and Office of Health Economics 2002 p38

[3] Culyer A.J. Introduction. in Towse, A., Pritchard, C. Devlin, N.(Eds) Cost effectiveness thresholds: Economic and Ethical Issues. London Kings Fund and Office of Health Economics 2002 pp9-15

[4] Rawlins MD. Culyer AJ National Institute for Clinical Excellence and its value judgements. BMJ 2004;329:224-227

[5] Devlin, N. Parkin, D. Does Nice have a cost effectiveness threshold and what other factors influence its decisions. Health Economics 2004; 13:437-452

[6] Culyer AJ., McCabe CJ., Briggs, AH., Claxton, K., Buxton, MJ., Akehurst, R., Sculpher, MJ., Brazier, JE., Searching for a threshold not setting one: the role of the National Institute for Health and Clinical Excellence. J. Health Services Research and Policy 2007 12;1:56-58

[7] Williams, A. What could be nicer than NICE? London Office of Health Economics 2004

[8] Loomes, G. Valuing life years and QALYs: transferability and convertibility of values across the UK Public Sector in Towse, A., Pritchard, C. Devlin, N. Cost effectiveness thresholds: Economic and Ethical Issues. London Kings Fund and Office of Health Economics 2002 pp46-55

[9] Hutton, J. Maynard, A. A NICE challenge for Health Economics. Health Economics 2000;8:89-93

[10] Devlin N. An introduction to the use of cost effectiveness issues in decision making: what are the issues? in Towse, A., Pritchard, C. Devlin, N. Cost effectiveness thresholds: Economic and Ethical Issues. London Kings Fund and Office of Health Economics 2002 pp16-25

[11] Birch, S. Gafni, A. Cost effectiveness/utility analyses: do current decision rules lead us to where we want to be? Journal of Health Economics 1992; 11:279-296

[12] Gafni, A., Birch, S. Guidelines for the adoption of new technologies: a prescription for uncontrolled growth in expenditures and how to avoid the problem. CMAJ 1993;148:913-917

[13] Birch S. Gafni, A. On being NICE in the UK:guidelines for technology appraisal for the NHS in England and Wales.

[14] Sendi, P.. Gafni, A. Birch, S. Opportunity costs and uncertainty in the economic evaluation of health care interventions. Health Economics 2002;11:23-31

[15] NICE Guide to the Methods of Health Technology Appraisal London NICE 2004

[16] Claxton K., Value, Price Guidance and Evidence. Presentation to OFT London March 2007

[17] Personal communication Kalipso Chalkidou, NICE July 2007

[18] Faculty of Public Health Specialist Prioritising Health Services Conference, Birmingham 27th April 2007.

[19] Martin, S., Rice, N., and Smith, P. The link between health care spending and outcome. Evidence from English programme data. The Health Foundation London June 2007. http://www.health.org.uk/publications/research_reports/the_link_between.html (accessed 30 July 2007)

[20] www.nice.nhs.uk

[21] Association of Directors of Public Health. http://www.adph.org.uk/jan07.pdf (accessed 22nd August 2007).

[22] Rawlins M. Pharm J. May 2007.

[23] UK Office of Fair Trading. The Pharmaceutical Pricing and Reimbursement Scheme: a market study. London OFT February 2007.

[24] Seshami M. Gray, A. Time to death and health expenditure: an improved modelfor the impact of demographic change on health care costs. Age and Ageing 2004; 33:556-561

[25] Brazier, JE., Akehurst, RL., Brennan, A., Dolan, P., Claxton, K., McCabe, C., O'Hagan, A., Sculpher, M., Tsuchiya, A., Should patients have a greater role in valuing health states? Applied Health Economics and Health Policy 2005; 4(4):201-208