

INCORPORATING MULTIPLE CRITERIA IN HTA

METHODS AND PROCESSES

NANCY J DEVLIN AND JON SUSSEX

INCORPORATING MULTIPLE CRITERIA IN HTA

METHODS AND PROCESSES

NANCY J DEVLIN AND JON SUSSEX

MARCH 2011



Office of Health Economics
12 Whitehall London SW1A 2DY
www.ohe.org

CONTENTS

Executive Summary	1
1. Introduction	3
2. The inevitability of using criteria other than cost effectiveness in HTA	7
3. A continuum of processes for handling multiple criteria in HTA	13
4. MCDA methods	16
5. The use of MCDA in public sector decision-making	22
6. Conclusions	41
7. Options for taking MCDA forward	45

EXECUTIVE SUMMARY

Deciding what a health care system should pay for depends on the health gain that results, but not only on that. Other factors taken into account by the governments and insurers who fund health care systems typically include the impact on some or all of: social equity; the quality of the patient experience; impacts on the wider economy and the quality of evidence upon which to base a decision. There are multiple criteria for decision-makers to consider. This paper discusses how that might be done.

Health Technology Assessment (HTA) concerns the way that the costs and benefits of new health care technologies are identified and compared. The recommendations that result from HTA have important implications for patient access to new treatments and also affect the way resources are used across the health system for the benefit of all patients. Bodies undertaking HTA and using it to recommend what should be reimbursed by health care systems are increasingly common internationally. A leading exponent of HTA is the National Institute for Health and Clinical Excellence (NICE), which advises the National Health Service (NHS) at the national level in England and Wales. But HTA is increasingly relevant at the local level in the NHS, with the NHS reforms in England announced in July 2010 requiring GP-led consortia to decide on which health services to provide for the local populations they serve, and to be publicly answerable for those decisions.

Cost effectiveness is a key consideration in HTA and is a central feature of the approach to HTA taken by NICE, for example. However NICE like other HTA bodies around the world, emphasises that considerations other than cost effectiveness are also relevant to its decisions. However, despite the large quantity of documentation that NICE publishes on its website about the guidance it offers, it remains unclear to those outside NICE exactly how important the various criteria are relative to one another, and how they are taken into account in each decision that NICE makes about which health care technologies should be made available on the NHS.

This report considers a range of approaches that come under the general heading of Multi-Criteria Decision Analysis (MCDA). We explain what MCDA is and how it can be used. We show how MCDA is being used by other areas of the UK public sector, by local NHS organizations, and in health system decision-making in other countries. We consider the advantages of using MCDA, particularly in terms of greater transparency and consistency in decision-making. We also consider the costs and risks of MCDA, and how to strike the right balance between flexibility and consistency in decision-making.

We argue for greater use of MCDA as an aid to HTA based decision-making in the NHS. The same applies to other health care systems internationally. The discipline of explicit identification and weighting of the criteria upon which health care resource allocation decisions are made is valuable. MCDA also makes it easier to hold decision-makers to account for the decisions they make on behalf of the public, than when decisions are based on more opaque deliberative processes, and hence should lead to greater public confidence in the decisions that result.

EXECUTIVE SUMMARY

We end our report by outlining the research that could be undertaken to support that.

The purpose of this OHE report is to inform and stimulate debate about the way different sorts of evidence and considerations are taken into account in decisions about new health care technologies. The report will be of interest to anyone interested or involved in policy and practice of health technology appraisal – and, more generally, to anyone involved in resource allocation in health care.

1 INTRODUCTION

Health Technology Assessment (HTA) concerns the way that the costs and benefits of new health care technologies are identified and compared. The recommendations that result from HTA have important implications for patient access to new treatments and also affect the way resources are used across the health system for the benefit of all patients.

Bodies undertaking HTA and using it to recommend what should be reimbursed by health care systems are increasingly common internationally. In England and Wales, HTA for the National Health Service (NHS) is undertaken at the national level by the National Institute for Health and Clinical Excellence (NICE). NICE has become one of the foremost HTA bodies in the sophistication of its methods. In this report we use NICE as a leading example, but the discussion is equally relevant to HTA regardless of who undertakes it. Indeed, HTA is increasingly relevant at the local level in the NHS, with the NHS reforms in England announced in July 2010 requiring GP-led consortia to decide on which health services to provide for the local populations they serve, and to be publicly answerable for those decisions (Department of Health 2010).

In the appraisal of health care technologies, NICE and many other HTA bodies rely heavily on cost effectiveness evidence. Specifically, NICE considers evidence on the addition to cost per Quality Adjusted Life Year (QALY) gained by a technology, relative to an alternative comparator treatment, as its principal approach to gauging cost effectiveness (NICE 2008a). That is, it estimates the additional cost of using the new treatment and divides that by the additional QALYs it produces compared with the alternative treatment to give the ‘incremental cost effectiveness ratio (ICER)’. Whether any given ICER is or is not deemed to represent good value for money depends on its comparison against NICE’s cost effectiveness threshold, intended to represent the opportunity cost to the NHS.

However, while cost effectiveness is the dominant consideration, other factors are also considered relevant and are taken into account by HTA bodies including NICE. Indeed NICE is increasingly clear about what these factors are, and the way that it reflects these ‘social value judgements’ in its decisions (Rawlins et al. 2009; NICE 2008b).

Nevertheless, it remains unclear to those outside NICE exactly *how* important these other considerations are, and how they are incorporated into decision-making. The identification of these factors by NICE tells us that they must count for *something*, but not how much. That is, we do not know what weight is attached to each in the decision-making process, or the trade-offs that NICE is prepared to make between cost per QALY and these other factors. Furthermore, the information provided in published NICE guidance “may not fully reflect all of the individual factors considered by the Appraisal Committee at the time of the appraisal” (Tappenden et al. 2007).

Arguably, being more explicit about the factors that influence decisions, and the way these are taken into account, could serve to:

- Improve the transparency of the decision-making process and the accountability of NICE to taxpayers (Baltussen and Niessen 2006);
- Improve the consistency of decision-making – for example, by ensuring that each of NICE’s four appraisal committees treat these considerations in a similar manner;

1 INTRODUCTION

- Facilitate greater consistency between the way NICE decides on new technologies and the way the NHS decides how to allocate its budgets;
- Provide an opportunity for NICE to engage the public in decisions about what criteria to use, and their relative importance – leading to more ‘buy-in’ to the difficult decisions NICE is required to make;
- Sharpen the signals to industry about what aspects of innovation NICE (acting as an agent for the NHS) values and where research and development (R&D) efforts should be directed.

A wide variety of methods exist for taking multiple criteria into account in decision-making, including the deliberative decision-making approach currently employed by NICE (Culyer 2009), through to a broad set of methods and approaches which come under the heading of multi-criteria decision analysis (MCDA)¹.

We define MCDA in broad terms to be:

MCDA: A set of methods and approaches to aid decision-making, where decisions are based on more than one criterion, which make explicit the impact on the decision of all the criteria applied and the relative importance attached to them.

This definition of MCDA encompasses a wide range of different approaches, both ‘technical’ and ‘non-technical’ in nature. Some types of MCDA involve sophisticated algorithms to suggest optimal choices; others simply aim to provide some structure to the deliberative process. All aim to facilitate replicability and transparency in decision-making.

Replicability in HTA: The extent to which a different committee, or the same committee on a different occasion, presented with the same evidence and social value judgements, would make the same decision.

Transparency in HTA: The extent to which the evidence and other factors taken into account in decision-making, and the decision-making process itself, are evident to observers outside that process.

Thus MCDA techniques could add value to a wide range of decision-making tasks. But it is important to be clear that MCDA is a *support* to decision making; it is not a prescriptive tool, as it is sometimes misrepresented. All multi-criteria decisions require degrees of judgement: MCDA aids the exercise of that judgement.

MCDA methods are routinely used in other areas of the UK public sector (Puig-Peiro 2011). Examples may also be found of MCDA being used to make decisions about access to health care in other countries. However, NICE has thus far expressed a reluctance to move toward a more formalized, quantitative treatment of multiple criteria – a position which was reinforced by a 2009 review commissioned by NICE and conducted by Sir Ian Kennedy, who argued, with respect to MCDA, that:

“First, the claim of greater transparency is illusory. What it would introduce is some entirely mechanistic process, in which benefits are arbitrarily given the same weight inter se and over

¹ The literature also refers to these same approaches as Multi-Criteria Analysis (MCA) and Multi-Criteria Decision Making (MCDM). For consistency, we use the expression MCDA throughout.

1 INTRODUCTION

time. But, what is at stake is a set of judgements, which, while they must be based on clear criteria, should not be reduced to some exercise in spurious numerical certainty simply through adding them up. Indeed, there is no validated mechanism for doing so." (Kennedy Review 2009, p. 29 – emphasis added)

These statements suggest that the nature and role of MCDA in decision-making is somewhat misunderstood. Although it would be possible to apply MCDA in a way that is “entirely mechanistic” and that imposes equal weights on each criterion, that is wholly unnecessary and very likely to be unwise. MCDA does not imply either of those features. It is an aid to judgement, not a replacement for it. Furthermore there are several validated mechanisms for combining multiple criteria in a logical and explicit way, as we demonstrate later in this paper.

That same report goes on to argue that:

“Because I have concluded that those benefits which I say should be taken account of should (be – sic) incorporated into NICE’s estimation of health gains as against health losses, the appraisal system should make it clear how this is to be done... But it must do so in a way that does not perpetuate the unfortunate idea, which could currently be entertained, that there is a methodology based on ICER/QALY and then there is some set of afterthoughts. If indeed social judgements, values or benefits do form part of NICE’s appraisal as NICE claims and it is a “deliberative process”, then they should overtly be identified as part of that deliberative approach...” (Kennedy Review 2009 p. 29-30 – emphasis added)

This echoes Culyer (2009), who notes:

“I do not think NICE is very good at weighing qualitative factors explicitly. In fact NICE is not alone, none of us is very good at doing it! But NICE ought to be better than it is. Nor is it very good at explaining recommendations of technologies with ICERs above the £20k threshold... There is quite a lot of confusion outside NICE (and possibly within it) about the meaning of the threshold range of £20-30k” (p. 14)

We now explore the ways in which greater clarity and consistency in NICE decisions might be achieved. Specifically, the aims of this paper are to:

- (a) assess the potential for using MCDA as part of the HTA process used by NICE and similar organisations internationally;
- (b) review the range of methods and approaches available, and their existing use in public sector and health services decision-making;
- (c) assess the potential advantages of, and also the issues and risks potentially associated with, the use of MCDA by NICE and other HTA bodies; and
- (d) identify options for the next steps which might be taken to further explore the use of MCDA to support decision-making in HTA.

Changes and new challenges currently facing the NHS make it particularly timely to investigate ways of strengthening decision-making regarding both new and existing health care technologies and services. First, the global financial crisis of 2008 and the economic recession that it triggered

1 INTRODUCTION

are causing substantial pressure on NHS budgets. This in turn creates pressure to ensure that new medical technologies help the NHS to improve productivity and performance, and do not displace existing health care services which are better value for money.

Second, the recent change in Government has led to further reforms of the NHS (Department of Health 2010a). Central to these reforms is a strengthening of commissioning to drive innovation and quality improvement. The introduction of an NHS ‘outcomes framework’, building in part on the existing patient-reported outcome measures (PROMs) programme (Devlin and Appleby 2010), will be used to guide GP commissioning of services and as a basis for measuring NHS performance. At the same time, ‘value based pricing’ (VBP) will replace the current approach to the regulation of branded medicines’ pricing, by linking prices more explicitly to a yet to be defined composite measure of ‘value’ (Department of Health 2010b).

This new system-wide focus on the value for money of different ways of spending health care resources creates an opportunity to ‘join up’ processes for the way new technologies are assessed by NICE in terms of effectiveness, cost effectiveness and value (however defined), with the way that the existing services that they either complement or displace are assessed by the wider NHS (Devlin 2010). The gains from a unified approach to priority-setting in terms of allocative efficiency – that is, for making sure NHS resources are used to maximum benefit – could be substantial. Progress is already being made by specialist commissioning units in more explicit use of multiple criteria in decision-making. The challenge will be to develop a consensus *across* the NHS about which criteria matter, and their relative importance.

We begin in Section 2 by considering in more detail why the consideration of multiple criteria is important, and indeed unavoidable, in decision-making about health care technologies. In Section 3 we describe the continuum of processes, from *ad hoc*, implicit approaches through to deliberative and more formalised decision-making approaches, that are available for incorporating multiple criteria in health care decision-making, and we place NICE’s past and current approach within that continuum. Section 4 focuses on technical aspects of MCDA: we provide a simple explanation of the principal types of MCDA methods and the ways in which these differ. We explain what MCDA does and also what it does not do. Section 5 reviews real experience with the use of MCDA in public sector decision-making in the UK and internationally, including case studies of the use of specific MCDA tools to inform health care decisions in Israel and New Zealand, with illustrative examples of some of the ‘tools’ available to support those processes. We conclude in Section 6 by summarising the potential benefits and some risks of adopting MCDA, and in Section 7 by offering suggestions about the next steps that might be taken to develop and test an MCDA process for the NHS or any other health care system.

2 THE INEVITABILITY OF USING CRITERIA OTHER THAN COST EFFECTIVENESS IN HTA

Cost effectiveness is widely used as a criterion to guide decisions about what to spend money on (and what not to spend money on) in collectively funded health care systems. Simply put, cost effectiveness tells decision-makers about the cost per unit of benefit produced (or, expressed another way, the amount of benefit per pound spent). If we had this information on all the health services that are or could be funded, and allow some simplifying assumptions, it would be possible to allocate the overall NHS budget each year in such a way that the total benefits produced from the approximately £100billion of taxpayer funding are maximised.

In practice, this simple view of how cost effectiveness might in theory guide decision-making confronts some important real-life complications (Birch and Gafni 2002, 2006 and 2007). For example, as it is generally practised, it assumes constant returns to scale and scope of the programmes being considered. Further, in practice, we do not have information on the ICERs of all options, just of those new technologies being assessed. Decisions are made as a series of one-off choices, rather than as an overall exercise in ‘constrained optimisation’. To make decisions about new technologies requires the ICER in each case to be compared against some ‘benchmark’ – the cost effectiveness threshold – to determine whether it constitutes acceptable value for money. However, the choice of the threshold has tended to be based on poor or no underlying evidence, and instead represents the decision-maker’s ‘best guess’ about what it should be (Appleby et al. 2009). This means there is a risk that the threshold might not be a good reflection of the opportunity costs in the health sector of adopting new, cost-increasing technologies. This has led some to suggest that budget impact – information on the total cost to the health care system of each option – should be considered alongside ICER evidence as an additional criterion (see Cohen et al. 2008, Niezen et al. 2009).

There are other reasons why the ICER on its own is unlikely to be adequate as a criterion for decision-making. Its legitimacy rests on the denominator – QALYs gained – being a good indicator of the benefits of health care and the value placed on those benefits. The basis on which QALYs are estimated and used relies on all QALYs being valued equally. However, there is a growing literature that suggests society may place a higher value on the QALYs gained by certain types of patients (for example the young or those with dependents), or certain sorts of health problems (for example, those with severe poor health and those at the end of their life) (e.g. see reviews by Dolan et al. 2004 and Shah 2009).

Further, there may be aspects of benefit which are valued by society, beyond the improvements in length and health-related quality of life measured by QALYs. The theoretical foundations of cost effectiveness analysis in extra welfarism do not require the measurement of benefit to be restricted to QALYs (Morris et al 2007). Examples of other sorts of benefits that *might* be considered relevant include: patients’ return to work (improved productivity of the workforce; reduced reliance on benefits); improvements in subjective wellbeing (‘happiness’); increased satisfaction with aspects of the delivery of health care (‘process of care’ utility); and reductions in health inequalities between population sub-groups.

In recognition of these sorts of issues, and the values society places on various other factors and considerations, HTA bodies (including NICE) generally *do* consider other criteria alongside cost effectiveness in their decision-making processes. A recent review of the criteria used in HTA (Golan et al. 2010) covering 11 countries and the US state of Oregon, suggests that the criteria that are taken into account fall into three main groups: (a) need, appropriateness and clinical

2 THE INEVITABILITY OF USING CRITERIA OTHER THAN COST EFFECTIVENESS IN HTA

benefits; (b) efficiency, including cost effectiveness; and (c) quality, solidarity and other ethical or social values. Table 2.1 categorises the principal criteria that these authors report were evident in HTA processes internationally.

Table 2.1 Main criteria and ‘other’ considerations used internationally for prioritizing new health technologies

Principles of allocative justice	Criteria
Need	<ul style="list-style-type: none"> • General • Severity of the condition • Availability of alternatives
Appropriateness	<ul style="list-style-type: none"> • Efficacy and safety • Effectiveness
Clinical benefits	<ul style="list-style-type: none"> • General • Effect on mortality (life saving) • Effect on longevity • Effect on health related quality of life
Efficiency	<ul style="list-style-type: none"> • Cost effectiveness/benefit • Budgetary impact • Cost
Equality	<ul style="list-style-type: none"> • General • Accessibility to the service • Affordability to the individual
Solidarity	
Other ethical or social values	<ul style="list-style-type: none"> • Autonomy • Public health value • Impact on future generations
‘Other’ considerations:	
Quality of the clinical and economic evidence	
Other considerations not elsewhere classified	<ul style="list-style-type: none"> • Strategic issues • Consistency with previous decisions and precedents

Source: Golan et al. (2010)

NICE, for example, says that it recognises a number of criteria as relevant to its technology appraisals, and that it does so by applying ‘special weightings’ to these criteria when making judgements about cost effectiveness – see Box 2.1. The way in which these factors are taken into account is set out in NICE’s social value judgement document (NICE 2008b):

“Decisions about whether to recommend interventions should not be based on evidence of their relative costs and benefits alone. NICE must consider other factors when developing its guidance, including the need to distribute health resources in the fairest way within society as a whole.” (Principle 3 – NICE 2008b p.18)

Box 2.1 Special weightings applied by NICE in making judgments about cost effectiveness

NICE takes a number of factors into account – and these are “given special weighting when making judgements about cost effectiveness” (Rawlins et al. 2009). The factors noted by NICE, with the examples provided by Rawlins et al. (2009) of specific decisions where these factors were taken into account, are:

1. Severity of the underlying illness

More generous consideration is given to the acceptability of an ICER in serious conditions, reflecting society’s priorities.

Taken into account in decisions about: Riluzole (for MND); Trastuzumab (advanced breast cancer); Imatinib (chronic myeloid leukaemia; Imatinib (for gastrointestinal stromal tumour); Pemexetred (for malignant mesothelioma); Omalizumab (for severe asthma); Sunitinib (for advanced renal cancer); and Lenalidomide (for multiple myeloma).

2. End of life treatments

The public places special value on treatments that prolong life at the end of life, providing that life is of reasonable quality.

Taken into account in decisions about: Riluzole (for MND); Imatinib (for gastrointestinal stromal tumour); Pemexetred (for malignant mesothelioma); Sunitinib (for advanced renal cancer); and Lenalidomide (for multiple myeloma).

3. Stakeholder persuasion

Insights provided by stakeholders e.g. on the adequacy of the measures used in clinical trials in reflecting symptoms and quality of life.

Taken into account in decisions about: Riluzole (for MND); Ranizumab (age related macular degeneration); Omalizumab (for severe asthma); Sunitinib (for advanced renal cancer); Somatotropin (growth hormone deficiency); and Chronic subcutaneous insulin infusion (childhood type 1 diabetes).

4. Significant innovation

Some products may produce demonstrable and distinct benefits of a substantive nature, and which are not adequately captured in the quality of life measures.

Taken into account in decisions about: Trastuzumab (advanced breast cancer); Imatinib (chronic myeloid leukaemia; Imatinib (for gastrointestinal stromal tumour); Ranizumab (age related macular degeneration); Omalizumab (for severe asthma); Sunitinib (for advanced renal cancer); Somatotropin (growth hormone deficiency); and Lenalidomide (for multiple myeloma).

5. Disadvantaged populations

Special priority is given to improving the health of the most disadvantaged members of the population e.g. poorer people and ethnic minorities.

Taken into account in decisions about: Pemexetred (for malignant mesothelioma).

6. Children

Given methodological challenges in assessing quality of life in children, society would prefer to give ‘the benefit of the doubt’.

Taken into account in decisions about: Somatotropin (growth hormone deficiency); and Chronic subcutaneous insulin infusion (childhood type 1 diabetes).

2 THE INEVITABILITY OF USING CRITERIA OTHER THAN COST EFFECTIVENESS IN HTA

These other factors are taken into account by NICE as mitigating factors relative to the cost effectiveness threshold range of £20-30,000 per QALY gained. Specifically, the decision-making process by which the ICER and other factors are combined is described as follows:

“...interventions with an ICER of less than £20,000 per QALY gained are considered to be cost effective. Where advisory bodies consider that particular interventions with an ICER of less than £20,000 per QALY gained should not be provided by the NHS they should provide explicit reasons (for example that there are significant limitations to the generalisability of the evidence for effectiveness). Above a most plausible ICER of £20,000 per QALY gained, judgements about the acceptability of the intervention as an effective use of NHS resources will specifically take account of the following factors.

- *The degree of certainty around the ICER. In particular, advisory bodies will be more cautious about recommending a technology when they are less certain about the ICERs presented in the cost-effectiveness analysis.*
- *The presence of strong reasons indicating that the assessment of the change in the quality of life inadequately captured, and may therefore misrepresent, the health gain.*
- *When the intervention is an innovation that adds demonstrable and distinct substantial benefits that may not have been adequately captured in the measurement of health gain.*

As the ICER of an intervention increases in the £20,000 to £30,000 range, an advisory body’s judgement about its acceptability as an effective use of NHS resources should make explicit reference to the relevant factors considered above. Above a most plausible ICER of £30,000 per QALY gained, advisory bodies will need to make an increasingly stronger case for supporting the intervention as an effective use of NHS resources with respect to the factors considered above.” (NICE 2008b p.18-19).

The recognition that factors other than cost effectiveness matter, and the explicit identification of these other criteria, is a commendable step toward greater transparency in the HTA process. The introduction by NICE, in 2009, of an explicit process for considering technologies that benefit patients at the end of their lives represents a further helpful step toward the more explicit treatment of these other criteria (NICE 2009a).

However, what remains unclear is:

- (a) whether or not taking these special circumstances into account altered the decision outcome that would otherwise have resulted;
- (b) the weight attached to each factor, and how each is ‘traded-off’ against cost effectiveness;
- (c) when more than one of these special considerations apply, whether the effect on the decision is additive or multiplicative; and
- (d) whether the list in Box 2.1 is exhaustive, or whether there are other, special considerations taken into account in individual decisions and, if so, how much influence those have.

For example, a recent econometric analysis of NICE decisions suggests that cost effectiveness alone predicts most of NICE’s decisions. Other factors may well be being taken into account – this is suggested by an estimated threshold of £40,000 per QALY, in contrast to the ‘official’ threshold range of £20,000-£30,000. But the treatment of these other factors is possibly not

2 THE INEVITABILITY OF USING CRITERIA OTHER THAN COST EFFECTIVENESS IN HTA

systematic enough for them to emerge as statistically significant drivers of decisions (Devlin et al. 2011).

Also relevant to the question of which criteria decision-making should address is the relationship between the payer and the HTA body, i.e. in England and Wales between the NHS and NICE. One way of thinking about this is that NICE is acting as an ‘agent’ for the NHS (the ‘principal’). That is, the NHS has delegated its decision-making about health technologies to NICE, which acts on its behalf to assess complex evidence and produce guidance. How close NICE is to being a ‘perfect agent’ for the NHS in England could theoretically be judged by whether the decisions NICE makes are consistent with those that would have been made in the NHS (e.g. by commissioners) had they access to the same information and evidence that NICE does. Put another way, and drawing an analogy with other ‘principal-agent’ relationships, to be a perfect agent NICE would need to have complete information not just about the evidence regarding a given technology, but also about the factors that are of importance to the NHS. This suggests that, in considering the way NICE uses multiple criteria in HTA, a relevant consideration is whether the criteria that NICE uses, and the weights it places on them, are in keeping with what is considered relevant to decision-making more widely in the NHS

While NICE has a well-developed and systematic national-level approach to the assessment of evidence, this seems to be somewhat at odds with the way priorities are decided sub-nationally in the NHS. NHS decisions sub-nationally appear to put less weight than does NICE on cost per QALY (Appleby et al. 2009), although in part this may be because that evidence is often lacking for existing services. Moreover, it is evident that many important health policy initiatives, with significant implications for NHS resource use, are driven largely or wholly by considerations other than QALY gains. Examples include waiting time targets and elimination of mixed-sex hospital accommodation. An analysis of the Department of Health’s ‘impact assessments’ of its policies and initiatives reveals that a wide range of types of benefits other than health improvement exert an influence on decisions affecting how the NHS’s limited budget is spent (Praet et al. 2011). While there have been a considerable number of localised NHS efforts to develop priority-setting frameworks (see Mullen 2004) – explicitly accounting for multiple criteria, as we will go on to discuss in Section 5 – the NHS currently lacks any consistent national framework to guide decisions made locally.

It is *inevitable* that NICE decisions will be based on multiple criteria – they already are. The question is not whether multiple criteria should be taken into account in these decisions, but what those criteria should be and how best to incorporate them in decision-making.

The current approach revolves centrally around cost effectiveness, combining ‘hard evidence’ on ICERs with ‘soft evidence’ (other considerations and social value judgements) in the deliberative process. This risks a ‘horse and rabbit stew’ problem²: while many ethical and social value considerations may be relevant to a given decision, the subtle flavour of those dependent on more qualitative evidence may be overwhelmed by the strong taste of the quantitative evidence on cost effectiveness. Ley (2003) notes that in HTA, qualitative dimensions of health care are not considered to be ‘real’ evidence. Some criteria may therefore dominate (that is, carry disproportionate weight) not because that reflects social priorities or values but because they are more readily measured and the evidence may seem more compelling.

² This expression was coined by the economist EJ Mishan in his influential 1976 book, *Cost Benefit Analysis*. However, his concern was the opposite: he argued that in CBA, the ‘rabbit’ is the scientific part of the evidence, and the ‘horse’ the various ‘intangibles’ that are problematic to measure and value in CBA.

2 THE INEVITABILITY OF USING CRITERIA OTHER THAN COST EFFECTIVENESS IN HTA

Deliberative processes carry a risk of unintended inconsistency in the way qualitative judgements are made across conditions, technologies and patients. The complex nature of these decisions, and the criteria that are relevant to them, pose a considerable challenge:

“When confronted with such complex problems, policy makers tend to use intuitive or heuristic approaches to simplify complexity, and in the process, important information may be lost and priority setting is ad hoc...policy makers are not always well placed to make informed, well-thought choices involving trade-offs of societal values.” (Baltussen and Niessen 2006 p. 2-3).

The risk of unintended omission of relevant information and of inconsistency in the way judgements are made from decision to decision are well researched in the decision theory and behavioural economics literature with respect to decision-making by *individuals*. Arguably, the issues become even more complex when decisions are made by committees. This suggests a role for an approach to priority setting that is more explicit, that provides a more systematic way of taking into factors relevant to social welfare, and that offers greater transparency and accountability to taxpayers and patients.

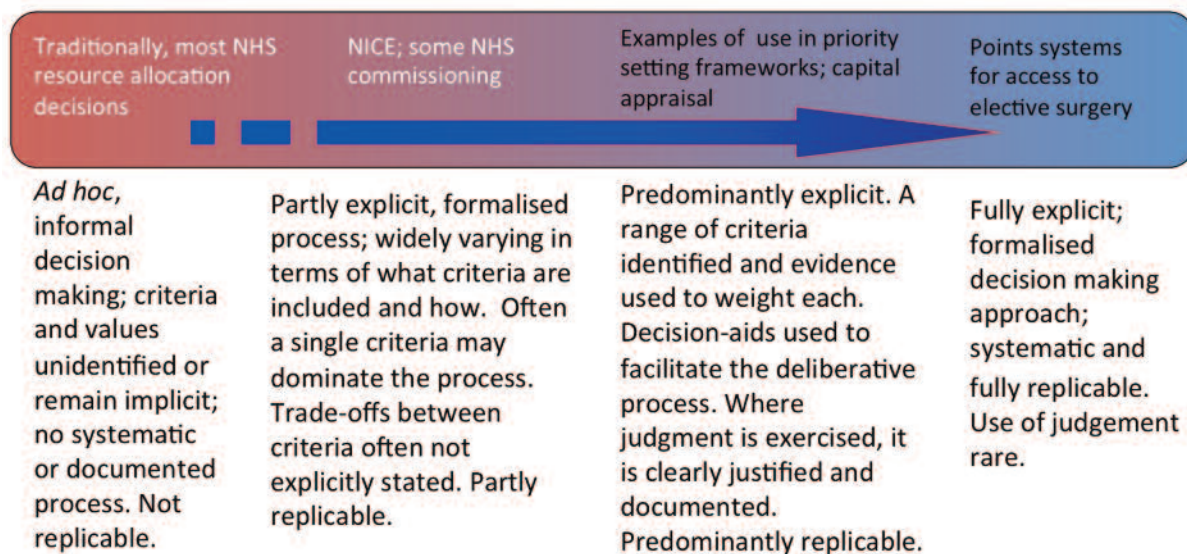
The set of techniques and methods that come under the umbrella of MCDA offers a pragmatic means of summarising the relevant information and *structuring* decisions and deliberative processes that involve multiple criteria. MCDA provides a systematic process for clarifying what is being taken into account (the ‘criteria’), how each of those criteria is to be measured, and how much importance (‘weight’) to put on each. In the next section we discuss the range of alternative processes available for coping with multiple criteria in HTA and resource allocation decisions.

3 A CONTINUUM OF PROCESSES FOR COPING WITH MULTIPLE CRITERIA IN HTA

Multiple criteria are encountered in many (arguably most) types of health care decisions, including those by individual patients and doctors regarding treatment choices, through to decisions made at the health system level that affect resource allocation and patient access to services.

Figure 3.1 presents a continuum of approaches to decisions involving multiple criteria. At one extreme, the relevant criteria may remain unidentified; the decision-making process is largely ad hoc; and the principles, values and processes that drive decision-making are not identified explicitly. Under these circumstances, decision-making is not transparent and is unlikely to be replicable. Historically, many resource allocation decisions made within the NHS were of this nature, with the principles and criteria of relevance to these decisions, and the nature of the trade-offs between the criteria, remaining largely implicit. There was lively debate in the early 1990s on the desirability of making ‘rationing’ decisions more explicit (see, for example, Doyal 1997 and New 1996). Greater explicitness may stimulate demand for better evidence, which leads to cost and delay in compiling that evidence, and a balance has to be struck between such costs/delays and the expected improvement to the quality of the decisions made. It may also suit decision-makers to keep their reasoning opaque in order to ward off unwelcome criticism.

Figure 3.1 A continuum of approaches to the existence of multiple criteria in health care decisions



NICE was established in England and Wales in 1999 and continues to be a leading example internationally of how health care decisions can be made in a more explicit manner. NICE is explicit about the criteria it uses but, as noted in Section 1, it is not fully explicit about the way it incorporates these in decisions. NICE’s efforts are particularly high profile, partly because it is a national level organization and also because of the controversy that attends decisions about access to new technologies for patients. However, for a number of years, local organisations in the NHS have also been making steady progress in developing and using explicit methods – prioritisation frameworks – to inform resource allocation decisions. In Section 5 we provide

3 A CONTINUUM OF PROCESSES FOR COPING WITH MULTIPLE CRITERIA IN HTA

examples of a number of these approaches: the way that they formulate the relevant criteria and the weights assigned to them are particularly interesting. Additionally, various regional specialist commissioning groups e.g. The London New Drugs Group, established by groups of Primary Care Trusts (the public bodies responsible for prioritising local health service provision in the NHS in England) to assist in decisions regarding specialist treatments, have adopted NICE-like processes.

There has therefore been a shift through time along this continuum by NHS organisations, with increasing use of explicit methods for addressing multiple criteria, but with considerable variation evident in the particular approaches being adopted.

At the other extreme of the continuum, ‘fully specified’ MCDA approaches are also possible. The use of ‘points systems’ to determine access to elective surgery is a rare example of the application of a fully algorithmic MCDA approach. These systems, which are used in Canada and New Zealand, assign points for each of a set of specified criteria (e.g. the clinical severity of a condition, its impact on quality of life, and the extent to which it would be ameliorated by treatment) and access to surgery is contingent on having a combined points score above a specified threshold. An important aim of such an approach is to ensure that decisions to refer and to treat patients are made in a systematic and consistent way with respect to the agreed criteria (see Devlin and Appleby 2010), so the use of *ex post* judgement in this decision context is much more constrained.

Where on the continuum should HTA be?

While it is clear that there is a continuum of approaches to handling multiple criteria, and that in general there is a shift toward the use of more explicit approaches, it is not obvious that the optimal approach to decision-making in all cases sits at the extreme right of this continuum. Rather, as we stated earlier, it is likely that some sorts of decisions will appropriately preserve greater degrees of judgement and flexibility than others, with MCDA used as a form of decision-support, rather than as a prescriptive tool.

What the ‘best’ MCDA approach is for HTA in part depends on the potential benefits from being more explicit in terms of greater transparency, consistency and replicability, as noted in Section 1. These need to be balanced against any costs of being ‘too explicit’. For example, the more formulaic the approach, the less scope there is to exercise judgement and to make exceptions. Where we are confident that the criteria used in HTA decision-making are based on clearly articulated principles; that the weights applied to them have a clear mandate in social or decision-makers’ preferences; and that those criteria and weights cover all issues of relevance in resource allocation decisions; then there may be little merit in allowing a role for *ad hoc* judgement. In practice, there may be a lack of clarity about the objectives of the health care system³ or the relevant principles that should determine the criteria. The more uncertainty there is over the appropriate criteria or weights, the more important the role of residual judgements and case-by-case reasoning. However, the discipline imposed by MCDA approaches when there is this lack of clarity – requiring careful thought about how to define all the criteria of relevance and how to trade them off, one with another – can aid the process of discovering the objectives and principles driving a decision.

³ In terms of the principal-agent model referred to earlier, it may be that the NHS (the principal) itself is unclear about the arguments in its utility function, or that there are considerable differences between local NHS organizations in this respect, making it difficult for NICE (the agent) to operationalise those on its behalf.

3 A CONTINUUM OF PROCESSES FOR COPING WITH MULTIPLE CRITERIA IN HTA

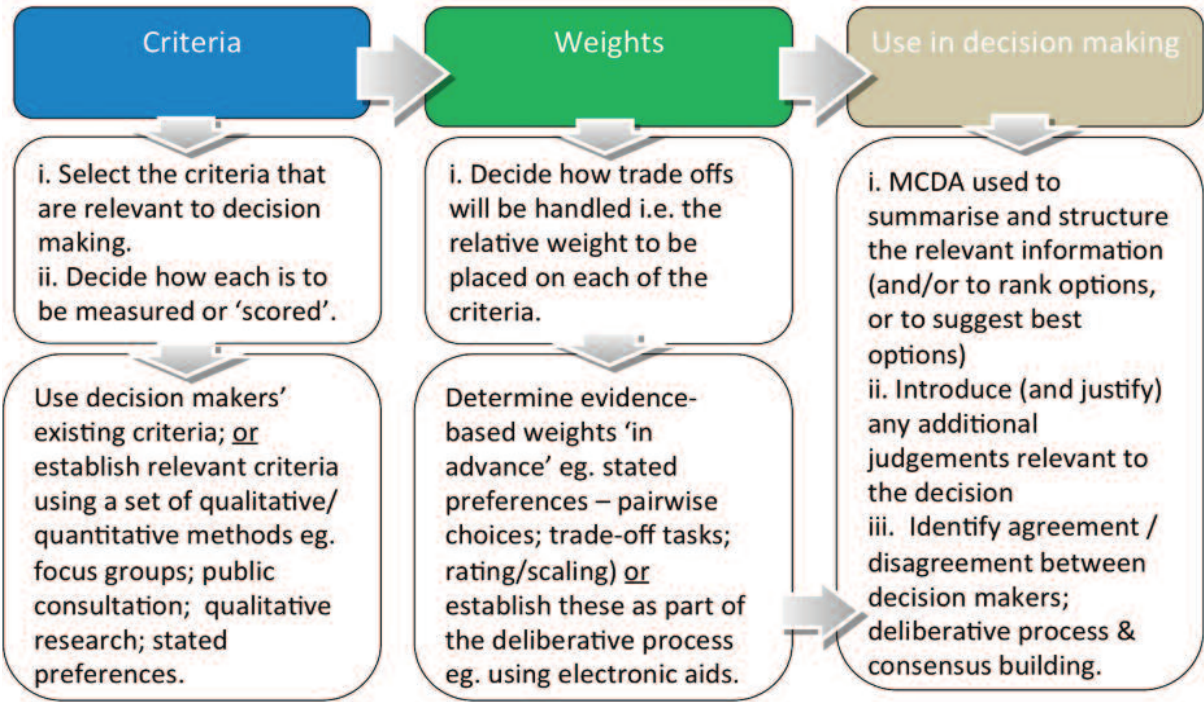
Another consideration is transactions costs: more explicit treatment of multiple criteria may not be costless. More explicitly acknowledging decision criteria may require the generation of additional evidence and data-gathering efforts. And a clearer process *may* (or may not) give rise to more legal challenges to decisions by stakeholders.

There is no ‘perfect’ HTA process. The issue is rather whether by adopting some form of MCDA approach NICE’s decision-making process, its decisions and, ultimately, the change in resource allocation that results, are better than the current process. In the next section we provide an overview of the range of MCDA methods that are available and the steps that would be involved in implementing them. In the light of that, and of the examples of MCDA use presented in Section 5, we will return in Section 6 to the issue of how to strike an appropriate balance between flexibility and consistency in decision-making.

4 MCDA METHODS

There are numerous different approaches to MCDA. All have in common an attempt to be clear about the criteria being taken into account, and about the influence of these multiple criteria on the decision. Beyond that, the methods and the way they are used in decision-making vary widely. An overview of the main elements is presented in Figure 4.1.

Figure 4.1 An overview of MCDA methods



In all cases, an MCDA requires the relevant criteria to be identified. The criteria might already exist and be part of a well-established decision-making process – for example, the criteria NICE uses, as set out in Box 2.1, could provide a ready-made basis for an MCDA. Alternatively, the relevant criteria could be established as a first step.

There are a number of means by which appropriate criteria might be suggested – and these might be used on their own or in combination. A desk-based search of the health economics, health services and bioethics literature would yield many practical examples of priority-setting frameworks (including those we have reported here), as well as the relevant underlying principles of procedural and distributive justice. If NICE’s processes are intended to reflect NHS priorities, then consistency with NHS priority setting frameworks or high level Department of Health objectives may be a key consideration. The views of the public and patients are also likely to be relevant, given the principles set out in the NHS Constitution (Littlejohns and Rawlins 2009). Examples of how the public might be involved are provided by the Oregon Health Services Commission’s efforts in developing a prioritized list of services to be funded under Medicaid (Blumstein 1997; Mitton et al. 2009) and New Zealand’s Health Funding Authority’s prioritisation framework (Ashton et al. 1999). In each case, the principles and relevant criteria were informed by extensive public consultation exercises. A further means of establishing the views of the general public is to employ qualitative research methods such as focus groups, or more structured approaches (such as Q methodology – see Baker et al. 2006) that combine qualitative and quantitative analyses.

4 MCDA METHODS

The criteria to be included in MCDA will need to have some characteristics. The criteria should be clearly defined and based on clearly articulated principles. It should be possible to measure or describe the characteristics of the options that decision-makers are considering in terms of those criteria. The criteria should be mutually exclusive (that is, they should not just be alternative measures or proxies of the same underlying principle), both to avoid double counting and because parsimony is desirable.

The decision about which criteria to include needs to be accompanied by a means of describing the options – for example, the health care technologies under consideration – in terms of those criteria. In some cases, criteria can be measured in their own units of measurement – for example, cost effectiveness could be included as incremental cost effectiveness ratios. In other cases, criteria may be captured as simple ordinal categories, e.g. the extent to which the technology addresses health inequalities might be measured as 0 (no relevance to health inequalities), 1 (some relevance in addressing health inequalities), or 2 (contributes substantially to addressing health inequalities). Where quantitative evidence exists on criteria, that may be represented either cardinally or by imposing ordinal categories – for example, cost effectiveness could be reported as 10 (< £20,000 per QALY), 5 (£20,000-£30,000) or 1 (>£30,000). What categories are used and how they are ‘scored’ require careful consideration.

Once an initial set of criteria has been identified, the weights to be applied to them need to be established. The weights can be established as an integral part of the MCDA itself, or alternatively a range of weights might be suggested by reference to other evidence, gathered via related studies or processes. In some cases the weights are applied to the scores (the numbers attached to the ordinal categories representing each criterion), to reflect the relative importance of each criterion in decision-making. In other cases, the scores used to describe each criterion might themselves be chosen to reflect their relative importance (bigger numbers denoting more importance) and then simply ‘added up’ across the criteria. Again, these represent important judgements and care needs to be taken to ensure that the scores and weights have a defensible basis⁴.

The relevant weights can be established as one product of an MCDA process – and there is a range of methods by which that might be achieved. For example, the weights might emerge from a deliberative process, where the evidence on each option, described in terms of each criterion, is considered. Some MCDA approaches, such as ‘decision conferencing’ (see Phillips 2006), entail methods for helping to structure those discussions, feeding back the decisions and implied weights via an iterative process. The outcome is a consensus on both the decisions themselves and the set of weights that have been applied. These in turn provide the framework for future decision-making processes.

Alternatively, an initial set of weights might be generated by a prior process designed to get selected participants (maybe the general public, or decision-makers themselves) to state their preferences. This draws on a set of well-established methods in economics (discrete choice experiments; conjoint analysis) to uncover participants’ preferences about the importance of the various attributes (criteria) through their choices. Participants are presented with a series of hypothetical options, each of which is described in terms of its characteristics on the relevant

⁴ In particular, simple additive points systems should be approached with caution, as the scores or points they employ are often arbitrarily selected rather than reflecting any underlying evidence.

4 PCT SURVEY ON CURRENT LEVEL OF SERVICE PROVISION FOR OBESITY

criteria. For example, one option may have an unfavourable cost effectiveness ratio but generate health improvements in those with very poor health; while another option might be highly cost effective at treating those with mild health problems. The choices that participants say they would make between these yield information about the relative importance they place on various levels of each criterion. These sorts of stated preferences exercises are widely used in health services research as a means of understanding preferences (for a review, see Ryan et al. 2008). Some MCDA tools incorporate these processes⁵.

These approaches to the determination of weights can be used on their own or in combination. The weights suggested by stated preferences tasks can provide a robust means of aggregating numerous multiple criteria into a more parsimonious set of criteria. And the weights attached to those criteria, elicited via stated preference tasks, can provide a starting point for discussions via a deliberative process.

Other, simpler approaches are also common. For example, some MCDA tools are simply handy, illustrative tools: one simply ‘plugs in’ a set of criteria, evidence on those criteria for the options under consideration, and a set of weights, and a preferred option is then suggested. Varying the criteria and the weights and looking at how this affects the suggested decision provides an easy way of illustrating and understanding the relationship between the value judgements embodied in the selection of criteria and weights and the suggested ‘best’ decision outcomes. Even very simple decision aids like this can be extremely helpful, especially in gaining a quick understanding of the robustness (or otherwise) of the suggested ‘best’ outcome to the selection of weights, in the absence of good evidence on what those weights should be. They also have the merit of being easy to understand.

However, simplicity is not always a virtue. Some rudimentary applications of MCDA involve describing each option under consideration on each of the relevant criteria by assigning arbitrary (i.e. arising from researcher value judgements, rather than with reference to any wider source of evidence on revealed or stated preferences) ‘points’ and then simply summing these up to arrive at an overall ‘score’. While such approaches may still be useful as part of a careful, deliberative process used to consider and recalibrate these weights, there is a risk that decision-makers are influenced by, or place too much credibility on, the initial, arbitrary scores that have been assigned⁶.

Testing the impact on the decision to be made of a variety of weightings across criteria is essential to using MCDA well. It is also something that MCDA tools make easy. Economists generally assume that the rates at which people trade off different attributes of services, including health care services, (i.e. the marginal rates of substitution between attributes) are not constant but can vary according to the quantities of each attribute that are on offer. Testing the impact of a range of weights, within bounds considered by the decision makers to be reasonable in the context of the particular decision to be made, rather than relying on a single set of weights, is highly desirable.

The range of possible MCDA approaches is wide but all approaches have important features in common, which are summarised in Box 4.1. It helps to be realistic about what MCDA does and

⁵ For example, the applications summarised in Boxes 5.7 and 5.8 later on do this using a method known as PAPRIKA (Potentially All Pairwise Rankings of All possible alternatives – see Hansen and Ombler 2009).

⁶ There is extensive evidence, in the stated preference literature, on the effect of ‘starting point bias’ and ‘framing effects’, which arguably might also be relevant to the way panels of decision-makers interact with information presented to them using some MCDA approaches.

does not do: see Box 4.1. It is worth stressing that no form of MCDA can avoid the need for value judgements; MCDA does not impose a fixed algorithm that allows no role for judgement.

Box 4.1 What does MCDA do? What does MCDA not do?

What does MCDA do?

1. Given a set of criteria, and weights for them, MCDA structures and combines that evidence to suggest a best choice, or a ranked list of options from best to worst, and can test the sensitivity of the choice or ranking to varying the weights for the criteria.

2. The preferred options identified by MCDA are likely to out-perform the use of intuitive judgement alone.

The decision theory and psychology literature abounds with examples of the various biases and heuristics that are evident when individuals are confronted with complex decisions (Kahneman 2003; Gilovich et al 2002). This is because the consideration of multiple criteria is cognitively demanding – arguably especially so when decisions are made in a committee context.

What doesn't MCDA do?

1. MCDA does not decide which criteria to include. That remains a matter for judgement.

There are various means by which that judgement might be reached. The list of 'other considerations' currently used by NICE provides a starting point. Another approach might be to engage in a consultative process with the general public, to obtain wider views on what other factors should be taken into account.

2. MCDA does not decide what weight to place on each criterion. That remains a matter for judgement.

Some MCDA tools (e.g. 1000Minds) incorporate into their approach a means by which the weights might be determined by asking samples of participants (who could be the decision-makers themselves or the general public) to make a series of pairwise choices, through which their preferences (i.e. the weight they attach to each criterion) can be discerned. Even that approach relies on a prior judgment about whose preferences and priorities should count – a non-trivial judgement in itself. Many MCDA tools do not incorporate that sort of preference-based approach, and rely instead on some other means of assigning weights to each factor. MCDA approaches that rely on arbitrary scoring and weighting (i.e. decided by the researchers rather than by others whose views are relevant) must be treated with extreme caution, as the weights may have little relationship with the relative importance society places on those factors.

3. MCDA does not replace decision-making – it facilitates it.

The purpose of MCDA is to clarify what choice would be made, if the criteria included are the only ones that matter, and if the weights applied to those capture consistent social preferences. But there may well be other, one-off considerations which are relevant to particular decisions. In these cases, decision-makers can judge such considerations to outweigh usual considerations. The use of MCDA will, however, require such departures to be carefully – and explicitly – justified.

The steps involved in using MCDA

MCDA is a well established and widely used approach to decision-making, and excellent user guides and manuals exist to support it. Of particular note is the extensive guide provided by the UK Department for Communities and Local Government, describing in detail the MCDA methods routinely used in that Department. Box 4.2, reproduced from that manual, provides a step by step account of the use of MCDA. In addition to our outline, above, of the ways that criteria may be selected, measured and weighted, Box 4.2 usefully emphasises the role of sensitivity analysis. Given that uncertainty may be a factor both in the evidence being considered, and in the MCDA process (e.g. it is more realistic to obtain ranges of plausible weights, than to pinpoint a precise weight), then exploring the sensitivity of suggested decision outcomes is a key part of the decision process.

Box 4.2 Applying MCDA – detailed steps

- 1. Establish the decision context.**
 - 1.1 Establish aims of the MCDA, and identify decision-makers and other key players.
 - 1.2 Design the socio-technical system for conducting the MCDA.
 - 1.3 Consider the context of the appraisal.
- 2. Identify the options to be appraised.**
- 3. Identify objectives and criteria.**
 - 3.1 Identify criteria for assessing the consequences of each option.
 - 3.2 Organise the criteria by clustering them under high-level and lower-level objectives in a hierarchy.
- 4. ‘Scoring’. Assess the expected performance of each option against the criteria. Then assess the value associated with the consequences of each option for each criterion.**
 - 4.1 Describe the consequences of the options.
 - 4.2 Score the options on the criteria.
 - 4.3 Check the consistency of the scores on each criterion.
- 5. ‘Weighting’. Assign weights for each of the criteria to reflect their relative importance to the decision.**
- 6. Combine the weights and scores for each option to derive an overall value.**
 - 6.1 Calculate overall weighted scores at each level in the hierarchy.
 - 6.2 Calculate overall weighted scores.
- 7. Examine the results.**
- 8. Sensitivity analysis.**
 - 8.1 Conduct a sensitivity analysis: do other preferences or weights affect the overall ordering of the options?
 - 8.2 Look at the advantage and disadvantages of selected options, and compare pairs of options.
 - 8.3 Create possible new options that might be better than those originally considered.
 - 8.4 Repeat the above steps until a ‘requisite’ model is obtained.

Source: UK Department for Communities and Local Government (2009).

The steps outlined in Box 4.2 have some similarities, in very broad terms, with the steps already involved in the process of health technology appraisal by organisations such as NICE (for example, see NICE 2009b). Both entail initial clarification of the context and options under consideration (for example, NICE’s “scoping” process); a review of the relevant evidence; and then some process of drawing that evidence together. MCDA would complement existing HTA processes by facilitating greater transparency both about what criteria are relevant, but more particularly about how the various types of evidence or considerations are combined, and trade-offs handled, to achieve a decision.

Principal classes of MCDA approaches

MCDA can be used to *describe* or present the options in terms of the performance of each option on each criterion (via a ‘performance matrix’, table or chart), with that information then being used by decision-makers to reach their own conclusions via a (qualitative) deliberative process.

Alternatively, MCDA can also combine and process the information. The outcomes of that process might be to suggest a single ‘best’ choice, or a ranking of options, or a set of preferred options given a budget constraint – depending on the number of options being appraised and the decision-making context.

The most obvious approach is simple linear aggregation, i.e. each score on each criterion is multiplied by the weight for that criterion and these weighted scores are then summed to determine an overall score for that option, which may be compared to the scores for other options under consideration. This is a very common approach in MCDA.

More sophisticated approaches use a range of methods from economics and operations research, including: Analytic Hierarchy Process (AHP); Multi-Attribute Utility (MAU) theory; Outranking and Data Envelopment Analysis (DEA). These each seek to establish the dominance⁷ or extended dominance of options, by drawing on various ways of establishing weights and combining scores across criteria. These same methods have common applications elsewhere in health economics. For example, MAU is the basis for research to estimate the ‘utilities’ (weights) for health states described by generic patient reported outcome measures, such as the EQ-5D, where those health states comprise different levels (scores) on different dimensions of health (‘criteria’). These utilities are, for example, commonly used in the calculation of QALYs in economic evaluation, where they are accepted as a legitimate basis for evidence and decision-making (NICE 2008). DEA is a technique commonly used to measure the efficiency of hospitals, where their output comprises multiple, dissimilar sorts of procedures (‘criteria’).

The various methods we have outlined here are perhaps more readily understood by way of example. In the following section, we provide a series of examples of the use of MCDA in public sector decision-making in the UK, including in the NHS, and in health care decision-making in other countries. The applications of MCDA are too numerous to provide an exhaustive account. Instead we have selected examples that either demonstrate a range of approaches or relate to decision contexts broadly relevant to HTA.

⁷ Dominance arises where one option is better than another on all criteria; or better on some criteria and no worse on others. Extended dominance involves processes for determining how options are to be ranked when they are better on some criteria but worse on others.

5 THE USE OF MCDA IN PUBLIC SECTOR DECISION-MAKING

Various forms of MCDA are widely proposed and frequently used in the UK public sector as tools to assist appraisal of options for spending public funds. The Treasury's "The Green Book – Appraisal and Evaluation in Central Government" (HM Treasury, 2003) suggests some form of 'multi-criteria analysis' be used where benefits or costs are not readily expressed in monetary terms – what it rather misleadingly refers to as "unvalued" costs and benefits:

"The most common technique used to compare both unvalued costs and benefits is weighting and scoring (sometimes called multi-criteria analysis). The basic approach to weighting and scoring involves assigning weights to criteria, and then scoring options in terms of how well they perform against those weighted criteria. The weighted scores are then summed, and these sums can be used to rank options. An even simpler method is to list the required performance criteria (sometimes called 'critical success factors'), and assess options in terms of whether they meet them or not." (HM Treasury, 2003, paragraph 5.78)

The Treasury guidance then refers for details to an earlier version of what is now the Department for Communities and Local Government's "Multi-Criteria Analysis – A Manual" (DCLG, 2009). The DCLG presents in detail a range of ways of conducting multi-criteria analysis, but without going so far as to require that particular approaches be used in specified circumstances. The Treasury guidance applies to all government expenditure other than that controlled by (elected) Local Authorities. The NHS falls within the definition of central government and the Treasury Green Book therefore applies to all NHS organisations.

But the Treasury's Green Book is explicitly a guide, not a set of mandatory rules, to allow for exceptions (but presumably not with the intention of permitting wholesale neglect).

Consequently practice around the UK public sector varies considerably. In this section we illustrate the range of ways in which MCDA approaches are being used in the NHS and in the UK public sector more widely, and highlight how appraisal criteria are chosen and weighted and how MCDA is used by decision-makers.

Use of MCDA in prioritisation by NHS commissioners

NICE is a national level prioritisation body for the NHS in England and Wales. But it is not the only one. The Advisory Group for National Specialised Services (AGNSS) makes recommendations for (or against) national commissioning of specialised services by the NHS in England for treatments involving no more than 500 patients (i.e. fewer than one per 100,000 population) annually across the country and/or is provided by no more than four centres (<http://www.specialisedservices.nhs.uk/pages/view/agnss>). AGNSS uses a decision-making framework containing 12 explicit criteria grouped under four headings (AGNSS 2010):

- Health gain:
 - Severity and ability of patients to benefit
 - Clinical safety and risk
 - Clinical effectiveness and potential for improving health
- Societal value:
 - Stimulating research and innovation
 - Needs of patients and society

5 THE USE OF MCDA IN PUBLIC SECTOR DECISION-MAKING

- Reasonable cost:
 - Average cost per patient
 - Overall cost impact and affordability including opportunity cost
 - Value for money compared to alternatives
- Best practice:
 - Best clinical practice in delivering the service
 - Economic efficiency of provision
 - Continuity of provision
 - Accessibility and balanced geographic distribution

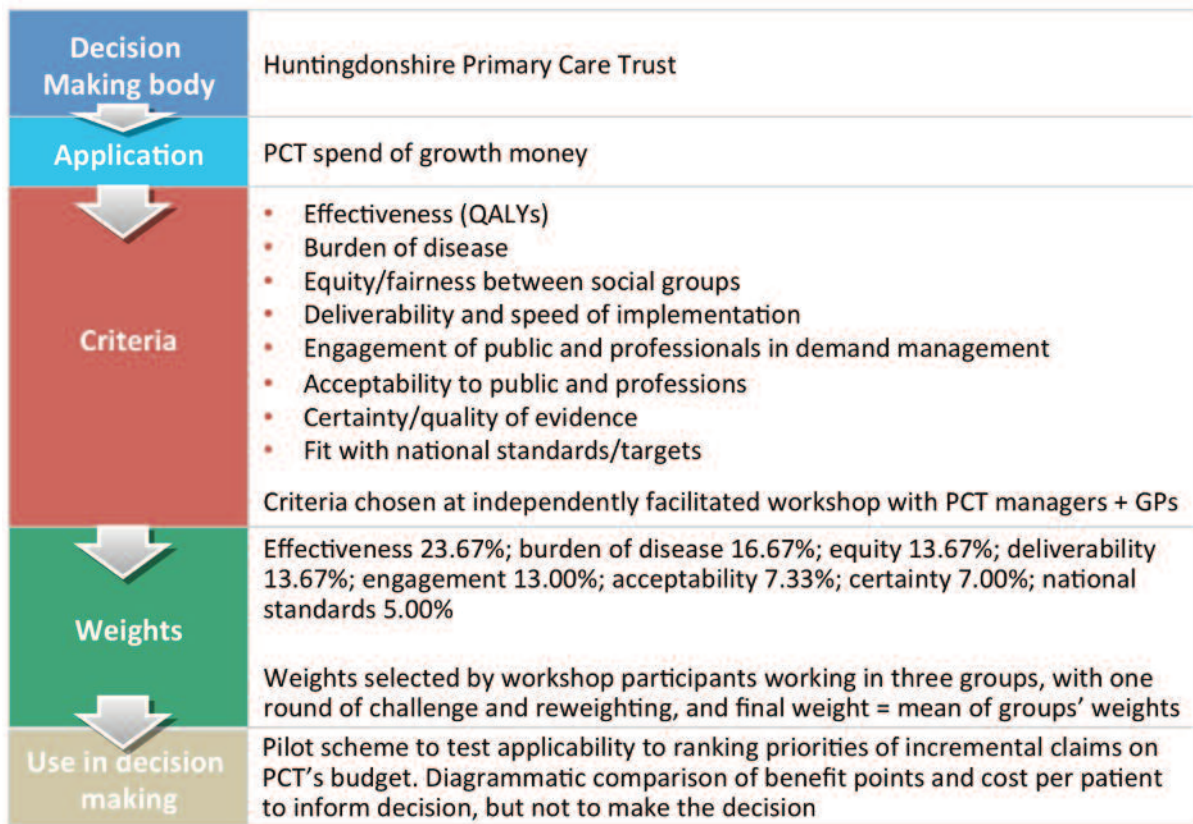
The list of criteria is explicit but no weights are discussed, i.e. no indication of their relative importance, nor any requirements that any of the criteria be quantified. The AGNSS website does not (yet) contain any examples of how these criteria have been used to make a specialised service commissioning decision.

The published literature contains references to a large number of examples of sub-national NHS commissioning bodies (Primary Care Trusts and their Health Authority predecessors in England, and their counterpart organisations in the NHS in Northern Ireland, Scotland and Wales) using MCDA techniques in exercises to prioritise their spending. Prioritisation is usually between options for incremental spending at the margin rather than considering the totality of the commissioning organisation's expenditure. The use of MCDA is a complement to, rather than a substitute for, other sorts of approaches to priority setting, such as Programme Budgeting Marginal Analysis (PBMA) (see Peacock et al. 2009). Mullen's survey of NHS commissioners (Mullen, 2004) found 46 such prioritisation projects using MCDA based approaches, and her literature review found references to many more. The number will only have grown since then.

For example, Wilson et al. (2007) report an exercise they helped Huntingdonshire PCT with in 2004. The PCT wished to establish an explicit and defensible method for prioritising new demands on its budget. It chose an MCDA approach in which a one-day workshop of PCT managers (from both medical and non-medical backgrounds) and local GPs was used to select the criteria against which expenditure options would be tested and to give weights to those criteria. The criteria were worked up in discussion in three sub-groups of participants and the resulting three lists were then combined and sorted in a plenary session to end up with seven criteria. Box 5.1 specifies these and summarises the other key points of this particular MCDA exercise.

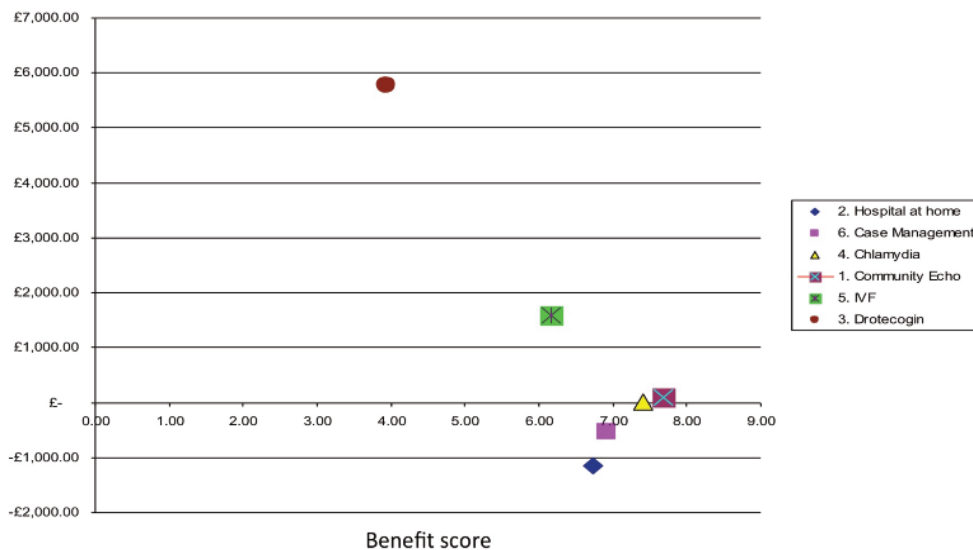
At a second one-day workshop six expenditure options were outlined in a concise briefing document (4-5 pages of text for each option, written within a common template) and then scored in terms of their abilities to satisfy those criteria. Scoring was again done in sub-groups who then reported back their scores in a plenary session. Where different groups had arrived at very different scores for the same option in terms of the same criterion, each group was asked to explain its rationale and was given the opportunity to modify their score. The overall score was taken as the average of the (modified where relevant) scores across the sub-groups of workshop participants. The total weighted score of each option, adding up across all criteria and using the weights shown in Box 1, was then calculated.

Box 5.1 Huntingdonshire PCT commissioning



These total scores were then compared with the estimated net cost per patient of implementing each option by plotting them on a graph as shown in Figure 5.1. The PCT used this diagram to inform its commissioning discussions but not automatically to make its expenditure decisions.

Figure 5.1 Scatter plot for six expenditure options showing mean benefit score and cost per patient



5

Source: Wilson et al. (2007), reproduced with the permission of the authors.

5 THE USE OF MCDA IN PUBLIC SECTOR DECISION-MAKING

A similar method, but with a more strategic focus, has been used by Isle of Wight PCT (Isle of Wight PCT, 2009; The Health Foundation, 2010). The PCT's Board identified five priority areas for commissioning over the coming five years. The PCT then ran six separate workshops "involving stakeholders and patients", which came up with 21 "strategic initiatives to improve quality of life and reduce health inequalities on the Island". These and the following two quotes are all taken from: Isle of Wight PCT, 2009:

"The issues and interventions emerging from the workshops were reviewed in a one-day Priority Setting Event with key stakeholders (executive and non-executive directors, patients' representative, Council representatives, GPs, Practice Based Commissioners, Professional Executive Committee and commissioning managers) to inform the strategic plan for 2008-13. The event was organized as a Decision Conference and facilitated by researchers from the London School of Economics.

In the meeting, participants built a model of the costs and value to the PCT of all twenty-one strategic interventions using the expert knowledge of workshop participants and the information provided by the Patients' and Stakeholders' workshops in a transparent, systematic process. Through facilitated discussion participants assessed the value of each intervention in terms of health benefit, reduction of health inequalities, probability of success and costs. The assessment consisted of scoring and weighting using Multi-Criteria Decision Analysis techniques."

Box 5.2 summarises the key points from this example of MCDA in action in the NHS. Four, rather high level, criteria were used. The available documentation does not show the relative weights given to them but the implication of the Commissioning Strategy document is that most weight was put on 'health benefit'. Each intervention was given a score for health benefit and to this was added some judgement of the impact on health inequalities between social groups.

The result was modified to take into account the probability of successful implementation, and was then compared with the cost of setting up and running the intervention over the next two years. It appears that a cost per benefit point scored calculation was then run and the 21 interventions ranked in ascending order of cost per benefit point. But the PCT is careful to stress that: *"The model, however, is by necessity a simplification of reality and was intended to inform the commissioning process and not to prescribe action."* (Isle of Wight PCT, 2009)

Box 5.2 Isle of Wight PCT commissioning strategy



There are many other examples of NHS commissioners using MCDA techniques to prioritise expenditure. Mullen (2004) reports that a major motivation for doing so is to “demonstrate openness, consistency and transparency in priority setting”. She also reports that NHS bodies that use MCDA vary widely in the specifics of their approaches. In particular, criteria are sometimes explicitly weighted, but sometimes not. Thus Mullen cites “Health Authority D” prioritising developments for use of additional funding over the next five years using seven criteria, each with their own explicit weight as follows and each scored on a 0-5 scale:

Criterion	Weight
Evidence of effectiveness	6
Value for money	4
Health gain or maintenance	4
Equity	2
Risk management	2
National or Board priority	1
Public preference	1

But even where weighting is not explicit, it unavoidably happens implicitly. For instance, simply adding up scores for an option against a number of different criteria implies that equal weight is implicitly being given to all the criteria, provided that the range of permissible scores is the same for each criterion. Scott and Lees (2007), cited by Mullen, refer to Argyll and Clyde Health

5 THE USE OF MCDA IN PUBLIC SECTOR DECISION-MAKING

Board in Scotland using a “Priority Scoring Index” with the following criteria and minimum and maximum permitted scores for each of them:

Criterion	Permitted range of scores
Potential health gain	-5 to +10
Prevention of ill health	0 to +5
Quality of life	-5 to +5
Equity of access	-5 to +5
Addressing health status inequalities at a population level	-5 to +5
Expressed demand	-5 to +5
Appropriateness	0 to +5
Strength of evidence	-5 to +5
Known priorities	0 to +5

Although no weighting of these nine criteria is referred to, the fact that “Potential health gain” has a range of possible scores with a width of 15 points (from -5 to +10) means that when the scores are added up this criterion has three times the impact of, for example, the criterion “Known priorities”, for which the range of possible scores is only 5 points wide (from 0 to 5). It is unclear whether this implicit weighting was deliberate and if so why disguising it within the scoring system was preferred to applying explicit weights of 3 to “Potential health gain” and 1 to “Known priorities” and having the range of possible scores the same for each criterion.

Both of the last two examples of MCDA in the NHS illustrate one of the potential pitfalls of this tool: namely overlapping or duplicative criteria. In the “Health Authority D” example, “Evidence of effectiveness” and “Health gain” look likely to overlap considerably. So, in the Argyll and Clyde example, do “Prevention of ill health” and more than one of the other criteria. Overlapping criteria mean that some aspects are being given heavier weight than may at first sight be apparent. It is preferable to separate out criteria into largely independent, discrete, categories so as to avoid that potential source of unintended consequences. Overlapping is more likely when long lists of criteria (e.g. 10 or more) are used and Mullen (2004) refers to reported difficulties experienced in practice in “weighting over-long lists of non-commensurate and overlapping criteria”.

Use of MCDA in NHS business cases for capital schemes

NHS organisations wishing to undertake major capital investments are required to obtain approval before committing resources to them. The approving body depends on the country and the scale of the proposed investment. In the NHS in England, projects with a capital cost greater than or equal to £75 million require approval by the Department of Health. In essence, approval is sought by submitting a business case which compares the proposed scheme against alternative options in a manner consistent with the Treasury Green Book guidance. The requirements are set out formally in the Department of Health’s “Business Case Guide” (Department of Health, 1994 – despite its vintage this is still the relevant guidance). The following paragraphs present two recent, publicly available examples that are currently (summer 2010) going through the approval process.

5 THE USE OF MCDA IN PUBLIC SECTOR DECISION-MAKING

The Epsom and St Helier University Hospitals NHS Trust in south east England is proposing a major redevelopment of one of its acute hospital sites: Epsom General Hospital. This is at the first stage of option appraisal, which is included within the ‘Strategic Outline Case, and consists of the sifting of a long list of options to a short list that will be appraised in more detail in a subsequent ‘Outline Business Case’. Department of Health, 1994, explains the required scope of these respective ‘Cases’ and of the later ‘Full Business Case’. The Strategic Outline Case for Epsom General Hospital (Epsom and St Helier University Hospitals Trust, 2009) describes how a long list of 13 options was reduced to a short list of six options via an MCDA process as follows.

Criteria were identified by managers at the Trust and then ratified with external stakeholders. It was unclear precisely who these external stakeholders were and whether they could add or remove criteria, but they are stated as including managers from the local Primary Care Trust (NHS Surrey). Fourteen different criteria are used and were attributed weights by the Trust who state that the preferences of all stakeholders were “taken into account”. Stakeholders were asked to rank which was the more important criterion in a series of pairs of criteria. The number of times each criterion was preferred was then summed to give a preference weight. The individual weights were then adjusted to sum to 100. Box 5.3 lists the criteria and weights. Each option on the long list was then scored for its achievement of each criterion, by the same group of Trust managers and external stakeholders. The range of possible scores appears to have been from 1 to 5 for each criterion, the higher the better.

Box 5.3 Strategic Outline Case for Epsom General Hospital redevelopment scheme

Decision Making body	Department of Health
Application	Strategic Outline Case for redevelopment of Epsom General Hospital i.e. sifting long list to short list for an option appraisal, for an NHS capital investment scheme
Criteria	<ul style="list-style-type: none"> • Patient safety (weight 12.8/100) • Meet quality standards (10.6) • Performance + outcomes (10.4) • Long term clinical and financial stability (9.5) • Productivity and efficiency of care services (7.9) • Patient focus (7.9) • Scope for modernisation and innovation (8.0) • Achievability (8.1) • Strategic fit (5.6) • Co-location of services (4.0) • Fits with “Centres of Clinical Excellence (4.1) • Fit with organisational cultures (4.0) • Stakeholders’ expectations (3.8) • Utilisation of estate (3.3) <p style="text-align: center;">Criteria selected by Trust managers and ratified with external stakeholders</p>
Weights	Weights, as above, were selected by Trust staff and (unspecified) external stakeholders
Use in decision making	Combined with scores (by Trust staff and external stakeholders) to give benefit points score. Capital cost per benefit point calculated but not used automatically. Trust proposes preferred short list options and seeks Department of Health and then HM Treasury approval

5 THE USE OF MCDA IN PUBLIC SECTOR DECISION-MAKING

The Trust then compared the benefit scores of each of the 13 options with the capital costs of building those schemes. Strangely they did not compare benefits with total costs, i.e. the combination of annuitized capital costs and annual operating costs. The latter would have been preferable, and operating costs are typically an order of magnitude greater than capital costs for NHS acute hospitals and are therefore more important to the ultimate choice of option, but perhaps operating costs were expected to be very similar, and therefore non-discriminating, across the options. The end result of the exercise was a table ranking the options in ascending order of capital cost per benefit point. This is reproduced as Figure 5.2 below (Epsom and St Helier University Hospitals Trust, 2009).

Figure 5.2 Table of capital costs and benefits points of options for Epsom General Hospital

Option	Description	Capital Cost	Benefit points	Cost/benefit point (£s)	Rank
1	Modern local Maximum Hybrid	£115,794,942	1,257.20	92,105	6
2	Modern local Intermediate Hybrid	£79,159,853	1,249.80	63,338	3
3	New concept Maximum Hybrid	£111,580,710	1,202.80	92,767	7
4	New concept Intermediate Hybrid	£76,925,977	1,195.50	64,346	4
5	Modern local Maximum New build	£384,538,768	1,183.10	325,026	12
6	Modern local Intermediate New build	£236,729,691	1,175.80	201,335	10
7	Major acute Maximum Hybrid	£131,014,653	1,163.00	139,920	9
8	Major acute Intermediate Hybrid	£99,831,794	1,155.70	105,893	8
9	New concept Maximum New build	£375,333,568	1,128.80	332,507	13
10	New concept Intermediate New build	£227,524,491	1,121.50	202,875	11
11	Modern local Maximum Refurbish	£91,622,794	1,111.30	82,446	5
12	Modern local Intermediate Refurbish	£66,202,175	1,103.90	59,971	2
13	Do Minimum: 21 st Century Turnkey Solution	£48,856,714	1,049.60	46,548	1

Source: Table 1 of Epsom and St Helier University Hospitals Trust, 2009. Reproduced by the permission of the Epsom and St Helier University Hospitals Trust.

5 THE USE OF MCDA IN PUBLIC SECTOR DECISION-MAKING

The Royal National Orthopaedic Hospital NHS Trust took a similar but simplified approach in its Outline Business Case for the redevelopment of its hospital (in Stanmore, north London). Nine criteria are listed and weighted out of 100 (Box 4), although who defined the criteria and attributed the weights, and how, is not stated in the Business Case document (Royal National Orthopaedic Hospital NHS Trust, 2010). This document describes just two options: the preferred redevelopment and a ‘do minimum’ option. Each was scored by a group of senior managers of the Trust and “a patient representative”. Detailed benefit points scores are reported and compared with the ‘equivalent annual costs’ of each option, i.e. the total of annual operating costs and annuitized capital costs over the expected life of the investment. The option with the lower equivalent annual cost per benefit point is preferred – i.e. the redevelopment option. Interestingly, that option dominates the ‘do minimum’ option, so there was no need to go as far as calculating the cost per point; i.e. redevelopment scored higher on every criterion than ‘do minimum’ and was also appraised to have lower equivalent annual costs. On this basis the decision is a ‘no-brainer’.

Box 5.4 Outline Business Case for the Royal National Orthopaedic Hospital, Stanmore, redevelopment scheme

Decision Making body	Department of Health
Application	Outline Business Case, Redevelopment of the Royal National Orthopaedic Hospital, Stanmore i.e. option appraisal, for an NHS capital investment scheme
Criteria	<ul style="list-style-type: none"> • Centre of excellence (weight 13.6/100) • Quality of clinical care (14.5) • Patient centred (8.2) • Access (9.0) • Elective capacity (11.9) • Workforce recruitment and retention (10.9) • Suitable for working with partners (10.8) • Deliverability/achievability (12.7) • Design, sustainability and quality of life (8.2) <p>Not stated how criteria were chosen</p>
Weights	Not stated how weights (as above) were attributed to criteria
Use in decision making	Combined with scores (by Trust staff and one ‘patient representative’) to give benefit points score. Set alongside equivalent annual cost (capital and running costs combined) but not used automatically. Trust proposes preferred option and seeks Department of Health and then HM Treasury approval

Use of MCDA by non-NHS UK public sector organisations

MCDA has been used for many years not only within the NHS but also by a wide variety of non-NHS public bodies, to assist them in making choices between competing options. It is impossible in a concise document to give more than a flavour of the range of uses to which MCDA is put. We present two very different examples here to demonstrate both the very different kinds of decision that can be aided by MCDA and the great importance and high profile of the resulting decisions.

5 THE USE OF MCDA IN PUBLIC SECTOR DECISION-MAKING

The UK Border Agency of the Home Office uses what it calls a “Points Based System” to assess applications by non-European Economic Area (non-EEA) immigrants seeking to work and settle in the UK. The rules are detailed and complex: the specifics vary according to the skill level of the applicant and whether sponsored by a UK employer. Box 5.5 summarises the approach for one major category of potential non-EEA immigrants to the UK: that for skilled workers sponsored by a UK employer (UK Border Agency, 2010).

Box 5.5 Applications to the UK Border Agency by non-EEA citizens to work and settle in the UK

Decision Making body	UK Border Agency of the Home Office
Application	Applications by non-EEA citizens to work and settle in the UK “Points Based System” for Immigrants Applying to Work in the UK Example: “Tier 2” applications, i.e. by skilled workers sponsored by a UK employer
Criteria	<p>“Sponsorship”, i.e. job offer from UK based employer</p> <ul style="list-style-type: none"> • Qualifications, i.e. level of highest obtained: PhD or Masters/Bachelors/A-level equivalent/below A-level • Prospective earnings • English language skills • “Maintenance”, i.e. ≥£800 of savings when you arrive to support self and dependents <p>How criteria were chosen by the Home Office is not stated</p>
Weights	<p>Explicit weights are used with maximum points for each criterion being:</p> <ul style="list-style-type: none"> • 50 for sponsorship • 15 for qualifications • 25 for prospective earnings • 10 for English language skills • 10 for maintenance (funds) <p>How weights were selected is not stated</p>
Use in decision making	To work in the UK, applicant must score a total ≥ 50 for the first 3 criteria, and 10 for each of the other two criteria. Insufficient points => application refused

Five criteria are deemed by the UK Border Agency to be relevant and the maximum possible points score varies from criterion to criterion (see Box 5.5), which has the effect of giving different weights to the different criteria. Being explicitly and fully sponsored by a UK-based employer is at first sight apparently the single most important criterion, worth 50 points at best, which is almost as much as the other four criteria put together as they offer a combined maximum of 60 points. But this version of MCDA is more subtle than calculating a single points total. Instead it requires 10 points to be scored on each of the last two criteria listed in Box 5 AND at least 50 points to be scored from the other three criteria combined (for which the theoretical maximum is 90 points taken together). Otherwise the application to settle and work will be refused. The UK Border Agency itself is the arbiter of how many points are scored by an applicant on each criterion.

Thus the UK Border Agency uses a variant of MCDA to make binary (yes or no) decisions that have a major impact on tens of thousands of individual applicants every year.

5 THE USE OF MCDA IN PUBLIC SECTOR DECISION-MAKING

The UK Department for Transport has long had a formalised approach to appraising and prioritising transport investment proposals. The current version has been named the “New Approach to Appraisal” (guidance to which is on the official website www.webTAG.org.uk). Five categories of benefit criteria are required to be taken into account: Environment; Safety; Economy; Accessibility (i.e. for users/consumers); and Integration (i.e. with other forms of transport, with the local economy and with other government policies).

The Highways Agency of the Department for Transport requires applications for road building and improvement schemes to be submitted with an “Appraisal Summary Table”. Box 5.6 summarises what this is required to contain.

Box 5.6 Appraisal Summary Table for a road improvement scheme

Decision Making body	Department for Transport
Application	Road improvement scheme Highways Agency, in line with Dept for Transport requirements for “Appraisal Summary Table” – E.g. A46 Newark to Widmerpool Improvement, June 2009
Criteria	<ul style="list-style-type: none"> • Environment: noise, local air quality, greenhouse gas emissions, landscape , townscape, heritage of historic resources, biodiversity, water environment, physical fitness, journey ambience • Safety: accidents, security • Economy: public accounts, economic efficiency of business users, economic efficiency of other users (“consumers”), reliability, wider economic impacts • Accessibility: transport option values, severance, access to transport system • Integration: transport interchange, land use policy, other Government policies <p>Criteria specified by DfT. How DfT chose is not specified</p>
Weights	Weights deliberately not used
Use in decision making	“To provide a systematic and consistent decision making basis to determine which trunk road schemes should proceed and the option to choose”

Numerous sub-categories of the five broad criteria dimensions are required to be evaluated for each proposed scheme, compared to the status quo option of not investing in the scheme. Quantitative measures have to be reported where available and monetised where possible but there is no calculation of an overall benefit value or benefit score. Instead the Appraisal Summary Table lists in some detail the nature of the expected impacts in terms of each criterion and sub-criterion, including the net present value cost. In the words of the relevant guidance document from the Department of Transport’s webTAG website:

“Those impacts that can be monetised are presented in monetary terms as well as in quantitative terms, but no other weighting information is provided. Decision takers must apply their judgement, taking account of the views of stakeholders determined through participation, to weigh up the impacts to reach an assessment of the overall value for money of the proposal.” (Department for Transport, 2004; paragraph 1.2.6)

5 THE USE OF MCDA IN PUBLIC SECTOR DECISION-MAKING

Figure 5.3 reproduces the pro forma Appraisal Summary Table for a road improvement scheme. Thus the Appraisal Summary Table is strictly a decision aid. It does not itself indicate whether or not the proposed investment is of high enough priority for resources to be committed to it. But at the same time, the Table is explicit about the criteria used and information upon which the decision is based. This permits public review of consistency between separate decisions, as well as the basis of each individual decision.

Figure 5.3 Department for Transport “Appraisal Summary Table”

Option		Description	Problems	Present Value of Costs to Public Accounts £m
OBJECTIVE	SUB-OBJECTIVE	(a) QUALITATIVE IMPACTS	(b) QUANTITATIVE ASSESSMENT	
ENVIRONMENT	Note			net properties win / lose NPV £m
	Local Air Quality			Concs wtd for exposure
	Greenhouse Gases			tonnes of CO ₂
	Landscape			Score
	Townscape			Score
	Heritage of Historic Resources			Score
	Biodiversity			Score
	Water Environment			Score
SAFETY	Physical Fitness			Score
	Journey Ambience			Score
	Accidents Security			PVB £m Score
ECONOMY	Public Accounts		Central Govt PVB, Local Govt PVB	PVC £m
	Business Users & Providers		Users PVB, Providers PVB, Other PVB	PVB £m
	Consumer Users			PVB £m
	Reliability			Score
	Wider Economic Impacts			Score
ACCESSIBILITY	Option values			PVB £m
	Severance			Score
INTEGRATION	Access to the Transport System			Score
	Transport Interchange			Score
	Land-Use Policy Other Government Policies			Score Score

Source: Department for Transport

www.dft.gov.uk/webtag/documents/project-manager/pdf/unit2.7.2.pdf

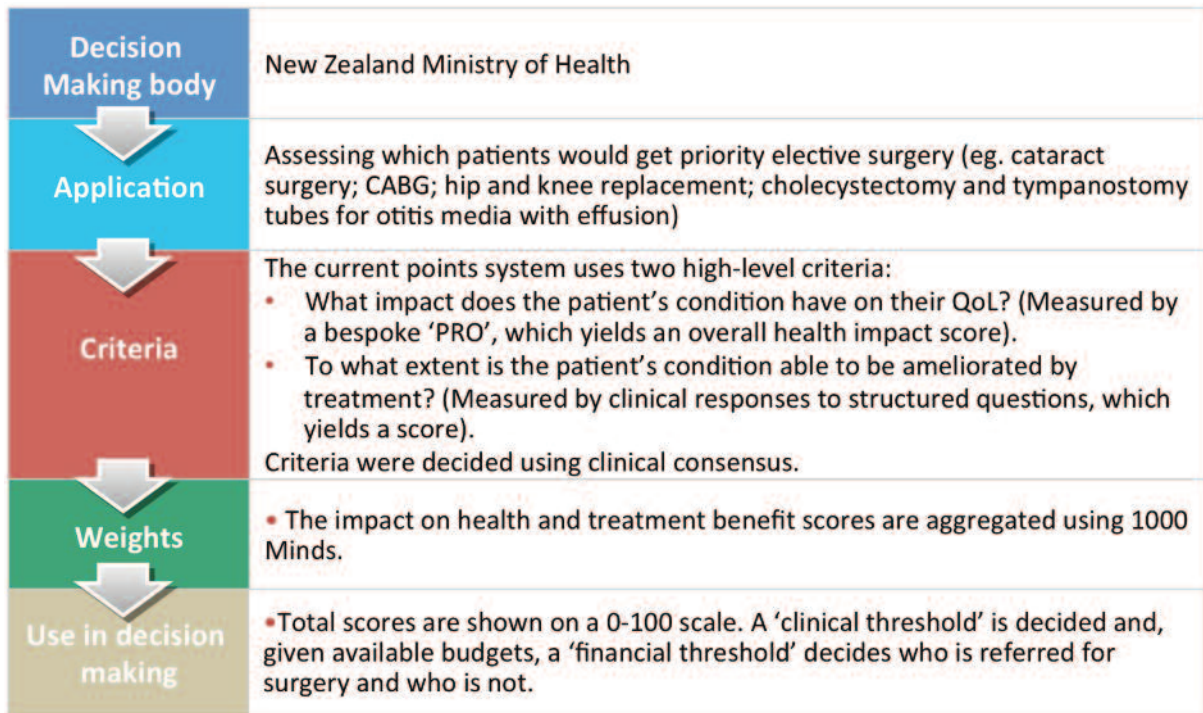
Use of MCDA in health sector decisions internationally

A wide range of examples of the development and application of MCDA in health care decision-making may be found in different decision contexts around the world. For example, Mussen et al. (2007) describe an MCDA approach for the weighing up of benefit and risk in regulatory approval of new medicines. Bots and Hulshof (2000) report and compare two cases of the use of a ‘participatory’ MCDA approach to priority setting in publicly-funded health care in The Netherlands. Nobre et al. (1999) report the use of MCDA to assist a Brazilian hospital’s decisions about the procurement of health technologies. Baltussen et al. (2006) describe the use of discrete choice experiments to derive the weights for a MCDA approach to priority setting in Ghana.

In this section we highlight two examples that illustrate quite different uses of MCDA. The first, New Zealand’s ‘points system’ for patient access to elective surgery, is notable because it provides a further example of a more ‘algorithmic’ use of MCDA in decision-making. The second, a use of MCDA currently being considered for pilot testing by the Israel Health Basket Committee, is of particular interest because of its direct relevance to HTA.

5 THE USE OF MCDA IN PUBLIC SECTOR DECISION-MAKING

Box 5.7 New Zealand's 'points system' for access to elective surgery



Naden and Barber (2010) describe the principles underpinning the explicit prioritisation of patients for elective surgery. The aim of this 'points system' is to ensure horizontal equity in access to publicly-funded health care services. That is, people of equal need and ability to benefit should have the same access to treatment regardless of their ethnicity, where they happen to live, or other factors. MCDA provides the basis for a points system which seeks to make these decisions about individual patients in a way that is explicit, fair and consistent. In doing so, it "marks a departure from a reliance on paternalistic, potentially idiosyncratic decisions by individual clinicians" (Naden and Barber 2010). Similar schemes are in place in Canada's social insurance system.

In New Zealand, points are assigned to individual patients based on two principal considerations:

- a) what impact does the patient's condition have on his or her quality of life? As this is best measured by the patient, this is measured by patient reported outcome measures;
- b) to what extent is the patient's condition able to be ameliorated by treatment? This is best judged by clinicians, and the points system provides a way of systematically recording that.

Crucial to the success of the points system has been the close involvement of clinicians in its development and in reaching a clinical consensus about the way patients are assigned points.

5 THE USE OF MCDA IN PUBLIC SECTOR DECISION-MAKING

This particular example of MCDA is of interest, because it has been in place (in various forms, employing different specific approaches to MCDA) for a number of years, so there is some evidence on its effects. The outcomes of the points system include:

- (a) important differences in access to surgery for people of equal need were revealed to exist by region, and were addressed;
- (b) important differences in access to surgery by people of equal need were revealed to exist by ethnicity and were addressed;
- (c) important disparities in access to surgery within clinical areas (e.g. cataract surgery vs. other treatments for vision problems), were revealed and this has informed resource reallocation within that clinical area; and
- (d) disparities between different areas of elective surgery, in terms of patients' relative ability to benefit, informs decisions about resource allocation between different clinical areas (Naden and Barber 2010).

A further example of the use of MCDA in health care is provided by Israel's 'Health Basket Committee'. This Committee has responsibility for considering all the new health care technologies that might potentially be added to the Israeli health care system's 'health basket', and allocating a fixed budget to these each year. For example, in 2008 a budget of 415 million Shekels (£63million) was available, from which a set of technologies (generally fewer than 100) must be selected from over 400 new technologies under consideration. Israel is currently considering piloting the use of MCDA (www.haaretz.com/print-edition/news/computer-may-prioritize-certain-health-basket-drugs-1.7008; and Golan and Hansen 2008). The criteria to be used in this pilot are described as follows:

- (a) benefit to Israel's population – e.g. in terms of 'lives saved', 'life prolongation', 'quality-of-life benefits', 'other social/ethical benefits' (e.g. targeted to children/minorities; reduces health gaps, and so on; i.e. 'equity' considerations). The technologies are ranked according to their 'total scores', from highest to lowest *Benefits*, represented on an index ranging from 100 (maximum benefit) to 0 (no benefit). In theory these *Benefit* 'components' could be expressed in terms of QALYs gained for the population overall;
- (b) net cost (e.g. annual) of the technology to Israel's health system – equal to the average net (incremental) cost per patient treated multiplied by the number of patients (over a given time period, e.g. one year);
- (c) quality of the evidence – especially concerning the *Benefits*, but conceivably also the Cost estimates (e.g. their robustness). Various *Benefits* 'grading' systems are available; e.g. the Oxford Centre for Evidence-based Medicine's *levels of evidence* (LOE): A, B, C, D;
- (d) any additional ethical or strategic x-factors or other considerations – i.e. other relevant special considerations not already included in (a) – (c) above.

Box 5.8 Israel’s Health Basket Committee: Pilot

Decision Making body	Israel’s Health Basket Committee
Application	Allocating a fixed budget (415m) across competing claims (>400) for new health care technologies to add to the ‘basket’ of those funded in the health care system (generally < 100 can be afforded).
Criteria	Multiple indicators of benefit were used: lives saved, life prolongation; quality of life benefits; availability of alternative treatments; other ethical/social benefits (eg reduces health gaps); aggregated into a single measure of Benefit (using 1000 Minds). Criteria: A Benefits; B Net Costs; C Quality of Evidence; D: other considerations
Weights	The Health Basket Committee is piloting using the VfM Charts (see Figure 5.4) as part of its deliberative process
Use in decision making	Using the VfM Chart, discussions proceed on ‘trial and error’ comparisons of alternative affordable combinations of technologies. The focus of attention begins at the chart’s top left quadrant, with possible tradeoffs in the NE direction (higher Net Cost, higher Benefit).

Source: Golan and Hansen (2008)

These criteria – assessed using available evidence, and weighted according to a preference elicitation procedure incorporated in 1000Minds (www.1000minds.com), are then reported back to aid decision-makers in a variety of ways, such as the bubble chart shown in Figure 5.4. Visual aids such as this are then used by decision-makers to assist their consideration of the fundamental question they face: given the available budget: which of these technologies should be selected?

MCDA tools being used in HTA

Both the applications of MCDA to health sector decision-making described in the previous section used a particular tool, 1000Minds. Many such tools exist, each having different properties and employing different underlying methods. In this section, we highlight two other tools that have been developed for, or are being applied to, HTA decisions.

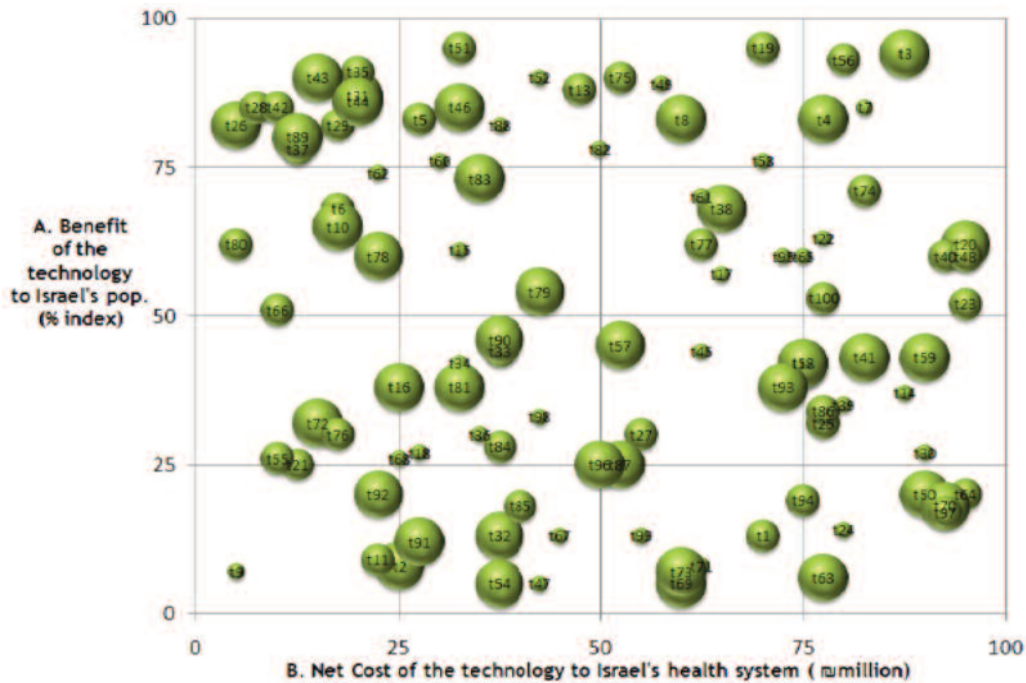
Café Annalisa

This MCDA tool has been developed by Professor Jack Dowie, a past member of a NICE Appraisal Committee. The applications of Café Annalisa are not restricted to HTA. For example, there are a number of applications of it as a decision aid for patients and clinicians (e.g. Masya et al. 2008 report its use in a study of treatment options for rectal cancer). However, the example of its being applied to NICE decision-making provided on the website (Café Annalisa 2010) provides a useful illustration of the way that such an approach *might* work.

5 THE USE OF MCDA IN PUBLIC SECTOR DECISION-MAKING

Figure 5.4 A new framework to help Israel’s Health Basket Committee: the VfM (value for money) chart

For 100 ‘illustrative’ (random data) new health technologies, labelled t1 – t100. Quality of evidence is represented by bubble size: large = high quality, medium = medium quality, small = low quality



Source: Golan and Hansen (2008)

The criteria included in the following example are:

Criteria	Explanation
pCIEff	Probability that that new technology is clinically effective
pCostEf 20k	Probability that the new technology is cost effective relative to the comparator at a willingness to pay of < 20k per QALY gained
Acceptability/Appropriateness/ Preferences (of public and patients)	(no additional definition provided)
Terminality	End of Life use
Orph/no alt/rescue	The new technology is an ‘orphan drug’ or it has no alternatives besides best supportive cae, or it is used in a ‘rule of rescue’ situation.
OtherEq	Other equity considerations
DH priorities	Clinical priority areas as designated by Secretary of State for Health and Welsh Assembly Government
HS Feasibility/Impact	(no additional definition provided)
Innovativeness	(no additional definition provided)
Wider Societal Considerations	(no additional definition provided)

5 THE USE OF MCDA IN PUBLIC SECTOR DECISION-MAKING

Figures 5.5 and 5.6 show an example of these criteria, applied to a hypothetical technology, relative to its comparator. The first part of the chart in each Figure shows the overall score, indicating which fares better, the new technology or the comparator. In Figure 5.5, the new technology is less attractive than its comparator: it has a lower score. That decision outcome is a product both of the characteristics and evidence on the new technology and of the weights attached to the criteria, in particular the dominance of cost effectiveness with a weight of 0.730.

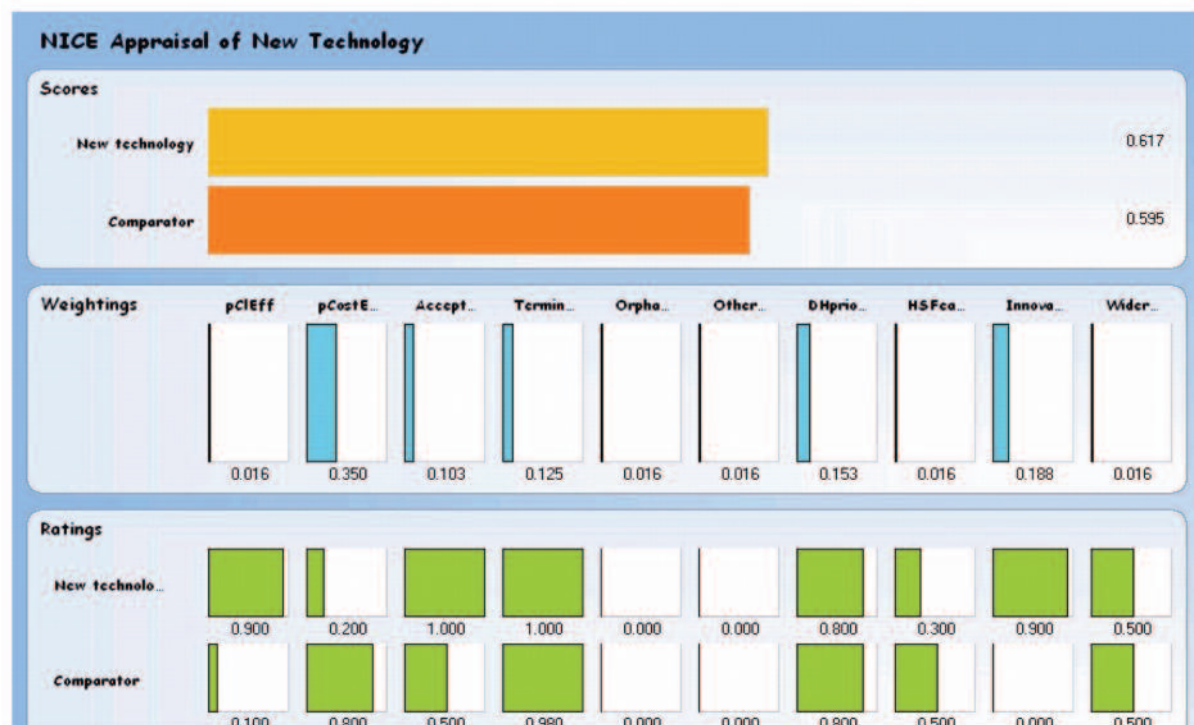
Figure 5.5 Illustrative application of MCDA to NICE decision-making: hypothetical example 1



Figure 5.6 shows that same technology and the same comparator and the same underlying evidence on each of the stated criteria as was the case in Figure 5.5. However, the weights attached to the criteria have changed: now the importance or weight attached to cost effectiveness has decreased to 0.350, with correspondingly higher weights attached to other criteria, such as 'terminality', 'DH priorities' and 'innovativeness'. With these different criteria weights the relative attractiveness of the options under consideration reverses and the new technology scores marginally higher than the comparator.

The purpose of this illustration is not to show that using other criteria would tend to favour new technologies; it is not clear that would be the case. However, what it does show is that MCDA provides a useful way of demonstrating the extent to which a given decision is contingent on (and sensitive to) the underlying weights that are used.

Figure 5.6 Illustrative application of MCDA to NICE decision-making: hypothetical example 2



Evidem

A relatively new HTA-focused MCDA approach is provided by EVIDEM (www.evidem.org), a collaboration of non-profit organisations established in 2009 to “promote public health through transparent and efficient healthcare decision-making via systematic assessment and dissemination of the evidence for and value of healthcare interventions”. EVIDEM provides a publicly-available decision-making framework and set of tools/instruments (Goetghebeur et al. (2008). There are some examples of applications of the MCDA approach it provides, e.g. Goetghebeur et al. (2010) applies it to the assessment of growth hormone for Turner syndrome patients.

The key elements of the EVIDEM approach comprise a standardised set of ‘intrinsic value criteria’ (“components that are quantifiable from a universal standpoint”) plus additional ‘extrinsic value components’ (components of decisions deemed not to be quantifiable from a universal standpoint, and which relate to specific health care system or ethical considerations).

5 THE USE OF MCDA IN PUBLIC SECTOR DECISION-MAKING

The intrinsic value criteria used in this tool are:

Intrinsic value criteria	
Disease impact	D1 Disease severity D2 Size of population
Context of intervention	C1 Clinical guidelines C2 Comparative interventions limitations (unmet needs)
Intervention outcomes	I1 Improvement of efficacy/effectiveness I2 Improvement of safety/tolerability I3: Improvement of patient reported outcomes
Type of benefit	T1 Public health interest (prevention; risk reduction) T2 Type of medical service (cure, symptom, relief)
Economics	E1 Budget impact on health plan (cost of intervention) E2 cost effectiveness of intervention E3 Impact on other spending (e.g. hospitalization, disability)
Quality of evidence	Q1 Adherence to requirements of decision-making body Q2 Completeness and consistency of reporting evidence Q3 Relevance and validity of evidence

While examples exist of their use, neither Annalisa or EVIDEM are currently embedded in any systematic, national or local level decision-making process.

6 CONCLUSIONS

Use of MCDA in decision-making

MCDA approaches, spanning a wide range of specific methods, are accepted as a useful means to assist decision-making in the public sector. As we have seen, MCDA is routinely used in some government departments in the UK. At a national level in the NHS, AGNSS has established an explicit list of criteria (but suggests no weights between them) for prioritising treatments for rare conditions, and there are many examples of various MCDA approaches being used to make local decisions in the NHS. We expect this practice to grow. For example, the New Cancer Drugs and Therapeutics Evaluation Committees in some regions of England (South East Coast and London) are establishing a scoring approach for prioritising the use of new cancer medicines. Regional specialist commissioners and the new GP Consortia require publicly-defensible decision-making frameworks. As both national and local commissioners confront the necessity to match investment in innovative new services with disinvestment elsewhere, MCDA-type approaches can provide a practical way of making those decisions in a systematic manner, given that those decisions will unavoidably require a weighing up of multiple considerations.

The development of MCDA frameworks offer an opportunity for dialogue between the NHS and NICE about the appropriate criteria and weights. One future possibility would be for NICE to assess and synthesise the evidence on new technologies, and for this to be reported to local GP-led commissioners via an agreed MCDA format. Local commissioners could then if they wished adjust the weights to account for agreed local preferences and priorities, tailoring the decisions to their own local circumstances. This might also provide a means of more directly taking local opportunity costs into account. As always, however, the greater freedom for local GP-led commissioners would come at the expense of national-level inconsistency in access to new technologies.

While we have not undertaken an exhaustive review of the applications of MCDA in decision-making⁸, of those examples we have provided in this report, there appear to be striking differences in the type of MCDA approach used in different public sector decision-making contexts. The two examples we found of ‘algorithmic’ MCDA were both about the ranking of individuals: the UK’s immigration points systems (also evident in other countries); and New Zealand’s points system for access to elective surgery. In contrast, decisions involving resource allocation (e.g. PCT commissioning) tended to use MCDA as a decision aid, to promote discussion and to structure the consideration of evidence. In these examples, the decision-makers’ use of MCDA was characterised by the discretion to exercise judgement, e.g. via the introduction of additional considerations, while being forced to be more explicit about where and why that judgement was used. The distinction between the approaches used in each of these contexts is interesting and as far as we know has not been explored in the literature.

One possible explanation for the more ‘mechanistic’ use of MCDA in making judgements about individuals is that considerations of fairness between individuals effectively preclude the use of ad hoc judgements as being unethical or difficult to defend. Departures from the criteria or weights to reflect individual circumstances would rapidly undermine the credibility of the decision-making process. Interpreted in another way, while there may be relevant individual circumstances, the number of individuals over whom decisions are made may mean that the transactions costs involved in considering exceptions while ensuring consistency between the treatment of individual cases is infeasibly high.

⁸ The use of MCDA in the private sector covers a vast territory we have not explored here at all.

6 CONCLUSIONS

In contrast, resource allocation decisions entail judgements between *groups* of people. In these decision-making contexts we found no examples of the prescriptive use of MCDA. Rather, a large element of judgement is exercised in the consideration of multiple criteria, with MCDA tools used to provide an intelligent way of structuring and testing those judgements, and of allowing greater consistency and transparency in reporting.

While judgement may be easier to exercise in resource allocation decisions between groups of people than in individual-level decisions, it is a matter for debate whether it is any more desirable. If the basis for selecting criteria and weights is clear, is supported by clearly articulated principles or agreed views, and provides an exhaustive account of the relevant considerations, then the rationale for allowing residual judgements is weak. However, if there is *uncertainty* over the criteria or weights, and if the *complexity* of decisions entailing a weighing up of the benefits between groups is such that all relevant factors cannot be incorporated into a parsimonious set of criteria, then judgement and case-by-case reasoning may be a necessary feature of the decision process and should be encouraged. The smaller number of decisions to be made (e.g. hundreds of new technologies vs. hundreds of thousands of individuals seeking to immigrate) also means that the transactions costs of basing decisions on deliberative processes are lower.

An alternative explanation might lie in public choice theory regarding interest groups (Stigler 1971). Because NICE decisions usually affect a known *group* (defined by their potential eligibility for a technology), the transactions costs of those individuals getting together to exercise collective influence is lower than the transactions costs of more dispersed, heterogeneous groups (e.g. those denied work permits). The other relevant group in this context is those (unknown) patients whose access to health services is adversely affected by reallocations of fixed budgets to accommodate new technologies. PCTs have an incentive to spread these opportunity costs over a wide range of programmes so the negative impact on any one patient is small. The affected group may be highly dispersed and difficult to identify: it is highly unlikely this group will exert collective power. Relying on deliberative processes and not being fully explicit about the decision process might in these circumstances provide the flexibility to favour groups with high collective interests.

Regardless of the explanation, it seems clear both that HTA will require the consideration of multiple criteria and that the process for doing so will retain a strong role for the exercise of judgement via deliberative processes. A pivotal question is how to strike an appropriate balance between ‘flexibility’ and ‘consistency’ in decision-making. This defines the potential role for, and gains from, MCDA.

Issues and risks

Our overview of MCDA methods in Section 4 suggests a number of important issues that would need to be addressed if MCDA is to be considered for use in HTA. There are also some risks.

MCDA entails a process for making decision criteria and weights more explicit. While well validated methods exist to support the development of MCDA, they all require a normative position (a value judgement) about *whose views should count*. One approach to applying MCDA

6 CONCLUSIONS

in HTA, as we discuss in Section 7, would be to start from the HTA body's existing criteria and to derive weightings from the deliberations that emerge from committee appraisal meetings. This would reflect a decision-makers' approach. Alternatively, both criteria and weights could be obtained by querying the general public.

In England and Wales, MCDA processes could provide a further means of supplementing the existing role of the NICE Citizen's Council by strengthening public involvement in NICE (Littlejohns and Rawlins 2009). Other possibilities include seeking patients' views, or those of other NHS decision-makers, or some mix of all of these options. The point is that this decision is a normative one that would need to be resolved before beginning work on an MCDA approach.

If MCDA builds on the HTA body's existing criteria – such as NICE's social value judgements – then the criteria themselves and the way that they are considered relative to the cost effectiveness threshold range would remain largely as it is at present, but with greater clarity around the weights attached to each criterion. However, if a more fundamental reform of the HTA process is considered, then this may have implications for the cost effectiveness threshold: both what that threshold *means* and *how it is used* would require a fundamental re-think.

Other risks relate to the reaction of patients and industry to the use of MCDA where it results in a decision not to recommend use of a technology. Would MCDA increase, decrease, or leave unchanged the probability of such decisions being challenged? At least one view has been expressed that improving the transparency of the process would increase appeals:

“As regards the claim that challenges would be fewer, it is at least arguable that such a system, so far from preventing challenges, would in fact promote them, as arguments over weighting and scoring went back and forth.” (Kennedy Review 2009, p. 29)

Depending on which MCDA approach is taken, the more explicit identification and incorporation of criteria other than effectiveness and cost effectiveness may generate demands for additional evidence and data collection and analysis. While many criteria may be simply categorical in nature (does the technology affect children?), criteria which seek to broaden the perspective (e.g. changes in productivity)⁹ would create an imperative to produce evidence to allow that to be measured. Providing evidence about a larger number of criteria adds to the costs of HTA, but does so to the same extent whether that evidence is used in an explicit MCDA process or a less explicit deliberative process. If the MCDA process requires information that does not yet exist about stakeholders' criteria and views on the relative weights of those criteria, then obtaining that information entails a cost. But it is a cost that a well-informed deliberative process would also need to incur.

The discussion in this paper has been focused on the potential to use some variant of MCDA within HTA processes, for example if it were to be adopted by NICE and other formal HTA bodies. In that context, MCDA could help to address concerns about the comprehensiveness, transparency and consistency of technology appraisal decisions on whether to reimburse a medicine (or other health technology), i.e. whether to permit patient access to it other than solely at their own expense, or on the maximum price that the health care system should be

⁹ The issue of the perspective from which HTA is undertaken continues to be contentious (Claxton et al. 2010). The arguments for and against MCDA are independent of the issues of the appropriate perspective; although MCDA would offer a pragmatic means by which broader 'societal' benefits might be incorporated.

6 CONCLUSIONS

willing to reimburse a medicine or other health care technology. Thus MCDA could, and in our view should, equally form part of the future “value-based pricing” of medicines promised for the UK in paragraph 3.23 of the 2010 Department of Health White Paper “Equity and excellence: liberating the NHS” (Department of Health 2010). The desirability of ensuring that assessment is comprehensive, transparent and consistent across decisions, is just as great for the value-based price of a new medicine (whichever public body is given responsibility for prices), as it is for decisions about access to those medicines at any given price. The more formulaic is the approach to MCDA, the greater is the transparency and ability to check the comprehensiveness and consistency of decisions made; but the less is the scope for negotiation about the implied price that results.

We are of the view that, overall, the advantages of MCDA processes in terms of greater explicitness about criteria, their relative weights and assessments of how well treatments achieve those criteria, and hence the greater accountability for the decisions made, make highly desirable their adoption nationally both for HTA and for medicines pricing purposes, and sub-nationally for the prioritisation decisions of local commissioners of health care. Deliberative processes that omit explicit MCDA may seem pragmatic but obscure the reasoning behind decisions of great interest to the public and at worst lead to unintended inconsistency between one decision and another, between reimbursing one treatment and refusing to reimburse another.

7 OPTIONS FOR TAKING MCDA FORWARD

Research can help to address some of the issues noted in the previous section, and would be an important aspect of the development of any MCDA approach. We outline in the following paragraphs three broad approaches that could be taken, ranging from an ‘incremental’ approach that builds on NICE’s existing criteria and processes through to the development of MCDA approaches that might entail more fundamental changes to the HTA framework.

Incremental approach

The purpose of this research would be to develop and evaluate a process for achieving greater transparency and replicability in decision-making – in effect, by shifting NICE just a little further along the continuum described in Figure 3.1. Taking as its starting point the existing approach to HTA, and working alongside NICE, the process could entail:

- (a) *ex ante* declarations (e.g. at the scoping stage) of which criteria are potentially relevant; and then
- (b) tracking the deliberative process and introducing additional structure to that by requiring the explicit recording of the judgments (weighing up of the evidence and other considerations) that are made about the importance attached to each criterion, and what those implied overall for any trade-offs against cost effectiveness.

Figure 7.1 illustrates the sort of simple recording template that could be used. This would address some concerns about the lack of transparency in the importance attached to these ‘other criteria’, i.e. those not captured in the ICER, while preserving the character of the NICE process by using both the existing criteria and the way of considering these as offsets against the cost effectiveness threshold range.

This approach builds on the more explicit reporting of decisions already possible under the End of Life guidance, which has allowed NICE to calculate the weight applied to the QALYs gained relative to the threshold in cases where that Guidance has been used (see NICE 2009a). Further, the information recorded in the templates could, considered over repeated decisions, be analysed to yield a set of indicative ranges of weights for each of the criteria, which could be built into future deliberations.

Figure 7.1 A template for explicit and transparent consideration of social value judgements in NICE’s deliberative process

To be considered at scoping:		To be considered at the appraisal committee:	
SVJ criteria	Relevant to this technology?	Record of committee’s deliberations on each SVJ deemed relevant at scoping: key points considered (free text)	Summary of the committee’s view of the importance of this SVJ in considering this technology: (1 = very important to 5 = not important)
End of life	Yes • No •		
Severity	Yes • No •		
Children	Yes • No •		
Social disadvantage	Yes • No •		
Small patient numbers	Yes • No •		
Lack of alternative treatments	Yes • No •		
Aspects of innovation not taken into account in the ICER	Yes • No •		
(other_____)	Yes • No •		
(other_____)	Yes • No •		
(other_____)	Yes • No •		
<p>Record of the overall (combined) impact of SVJs on the decision about this technology with respect to the cost effectiveness threshold range:</p> <p>Most plausible ICER for this technology £_____</p> <p>Implicit weight applied to QALYs gained from combined SVJs at £20k threshold*:_____</p> <p>Implicit weight applied to QALYs gained from combined SVJs at £30k threshold*: _____</p> <p>Summary of the overall influence of SVJs in the deliberative process for this technology:</p> <p>*“As the ICER of an intervention increases in the £20,000 to £30,000 range, an advisory body’s judgement about its acceptability as an effective use of NHS resources should make explicit reference to the relevant factors... Above a most plausible ICER of £30,000 per QALY gained, advisory bodies will need to make an increasingly stronger case for supporting the intervention as an effective use of NHS resources...”(NICE 2008, p.19).</p>			

‘Proof of concept’ MCDA research

Again, using NICE’s *existing* criteria as the starting point, research could be undertaken both to inform the range of weights suggested by social preferences and to explore the usefulness of using these criteria and weights to facilitate Appraisal Committee deliberations.

Plausible *ranges for* the weights to be applied to the criteria could be established via stated preferences research with samples of the general public. As we noted in Section 4, these methods are widely used in health services and health economics research. Values derived from stated preference methods with the general public are already accepted as having relevance to NICE decision-making: they provide the quality of life weights used to estimate QALYs. Using such methods to inform the selection of weights would provide a further means of involving and reflecting general public views into NICE decision-making.

An existing or bespoke MCDA approach could then be developed to incorporate those weights, and to explore the effect on decisions of the weights and the introduction of any additional one-off considerations considered relevant to a given technology. As we noted in Section 6, it is highly unlikely that a fully algorithmic MCDA approach would be either workable or considered desirable in HTA (or in resource allocation decisions generally). The MCDA approach that could be explored in research would reflect that, and be an aid for deliberative process.

Evaluating such an approach would require its use to be trialled in some way. One option would be to work in partnership with NICE both in the development of MCDA and in its use in deliberations. However, the introduction of a new decision-making approach in the context of a real technology appraisal carries considerable risks both to NICE and to its stakeholders. Given the gravity of NICE decisions, initial testing of MCDA decision aids might be better undertaken in a hypothetical decision-making scenario. For example, past (or present) Appraisal Committee members could be invited to participate in a shadow deliberative process, drawing on real evidence about real technologies (ones in which they had not previously been involved as individuals) for which previous decisions (recommended, not recommended, or ‘optimised’) are known.

Alternatively, the shadow appraisal committee could consider the same ‘real’ evidence as a ‘real’ appraisal committee, but working in parallel to it. However, that approach also carries some risks (for example, if the ‘shadow’ decision is inconsistent with the ‘real’ decision). A further option, which avoids the risks of contamination from knowledge of previous decisions, and the risks of running parallel processes, is to rely entirely on hypothetical evidence on hypothetical technologies as an initial step.

Evaluative research would need to clearly state, in the research design, the basis on which conclusions would be drawn about the MCDA approach. This requires careful thought. For example, this should include a clear means of judging the feasibility of the process; and the perceptions of Appraisal Committee members about its usefulness and legitimacy. Judging whether transparency has been improved – a key goal of MCDA – would require some means of seeking the views of patients, the public and industry on the way the decision and decision process are reported.

‘Clean slate’ MCDA research

A more ambitious approach might entail starting from a ‘clean slate’: undertaking research to establish both what a relevant set of *criteria* should be for HTA, and to establish the range of weights for those criteria.

Identifying the criteria could involve a mixed methods approach including both desk-based and primary research. For example, as we noted in Section 4, research methods such as qualitative analysis of focus group discussions and/or the use of more formal quantitative analyses of qualitative data (e.g. Q methodology) could be used to establish the relevant criteria.

The weights for those criteria could then be established using stated preference methods with samples of the general public or decision-makers. An interesting example of the kind of approach that is possible is provided by Tappenden et al. (2007), who report results from a discrete choice experiment with members of NICE Appraisal Committees. Committee members were presented with hypothetical scenarios described in terms of the levels on each of the criteria shown in Table 7.2 below, and were asked to indicate whether or not they would recommend the technology. While there are some limitations to that study, it illustrates the potential of using these sorts of techniques to suggest weights. Weights obtained using these experimental approaches could form a starting point for deliberations.

Table 7.2 Attributes (criteria) and levels (the way that criteria are measured) included in a discrete choice experiment with NICE appraisal committee members

Attribute	Description	Levels
Incremental cost effectiveness	Central estimate of cost effectiveness for the intervention compared to current standard treatment	0: £15k per QALY gained 1: £25k per QALY gained 2: £35k per QALY gained
Uncertainty	The degree of uncertainty surrounding incremental costs and effects	Low degree of uncertainty High degree of uncertainty
Age	The mean age of the population who will benefit from the intervention	0: Children (<18yrs) 1: Working (18-64 yrs) 2: Retired (> 64 yrs)
Baseline HR-QoL	An index utility score of patients prior to receiving the intervention, whereby ‘1’ represents a state of perfect health and 0 represents dead	0: 0.25 1: 0.50 2: 0.75
Availability of other therapies	Whether alternative effective therapies are available to manage the condition or not	0: No 1: Yes

Source: Adapted from Tappenden et al. 2007.

7 OPTIONS FOR TAKING MCDA FORWARD

As noted above, as important as the development of the MCDA processes themselves would be the design of the research required to evaluate them and the basis upon which valid conclusions might be drawn about the relative merits of the approach relative to current HTA processes. The 'clean slate' approach marks a more radical departure from current NICE processes, and would require a substantial programme of research, as well as posing additional practical challenges for trialling and evaluating the approach relative to current processes.

Concluding remarks

The gains from these research efforts could be substantial. Ultimately, however, those benefits will be realised only if the MCDA processes that are developed are used in practice. This in turn depends on the decision-making processes they deliver being credible, practical and relevant to decision-makers. While independent research efforts may yield interesting and innovative methods, research undertaken in partnership with decision-makers, and which is sensitive to the reality of the decision-making context, is considerably more likely to yield practical improvements that can be implemented.

REFERENCES

- Advisory Group for National Specialised Services. (2010) Decision-making framework for making recommendations on national commissioning. http://www.specialisedservices.nhs.uk/library/27/Decision_Making_Framework.pdf [Accessed November 19 2010].
- Appleby J, Devlin N, Parkin D, Buxton M, Chalkidou K. (2009) Searching for cost effectiveness thresholds in the NHS. *Soc Sci Med* 91 (3): 239-245.
- Ashton T, Cumming J, Devlin N. (1999) *Prioritising health and disability support services: principles, processes and problems*. A report to the National Health Committee on the Health Funding Authority's prioritization process. Wellington: National Health Committee [www.nhc.health.govt.nz/moh.nsf/pagescm/719/\\$File/prior.pdf](http://www.nhc.health.govt.nz/moh.nsf/pagescm/719/$File/prior.pdf) [Accessed August 10 2010].
- Baker R, Thompson C, Mannion R. (2006) Q methodology in health economics. *Journal of Health Services Research and Policy* 11(1):38-45.
- Baltussen R, Niessen L. (2006) Priority setting of health interventions: the need for multi-criteria decision analysis. *Cost effectiveness and resource allocation* 4: 14.
- Baltussen R, Stolk E, Chisholm D, Aikins M. (2006) Toward a multi-criteria approach for priority setting: an application in Ghana. *Health Economics* 15: 689-696.
- Birch S, Gafni A. (2002) Cost effectiveness/utility analysis. Do current decision rules lead us to where we want to be? *Journal of Health Economics* 11(3): 279-296.
- Birch S, Gafni A. (2006) Decision rules in economic evaluation. Chapter 46 in: Jones A (ed) *The Elgar Companion in Health Economics*. Edward Elgar.
- Birch S, Gafni A (2007). Being naughty about NICE? Questioning the methods used to maximize health gains from NHS resources. *Health Economics, Policy and Law* 2: 217-221.
- Blumstein JF. (1997) The Oregon experiment: the role of cost benefit analysis in the allocation of Medicaid funds. *Soc Sci Med* 45 (4): 545-554.
- Bots P, Hulshof J. (2000) Designing multi-criteria analysis processes for priority setting in health policy. *Journal of Multi Criteria Decision Analysis* 9: 56-75.
- Café Annalisa. (2010) www.cafeannalisa.org.uk/STARTERS/Decision/NICE_Appraisals/ [Accessed June 15 2010]
- Claxton K, Walker S, Palmer S, Sculpher M. (2010) Appropriate perspectives for health care decisions. CHE Research Paper 54. Centre for Health Economics, University of York. www.york.ac.uk/media/che/documents/papers/researchpapers/CHE%20Research%20Paper%2054.pdf
- Cohen J, Stolk E, Niezen M (2008) Role of budget impact in drug reimbursement decisions. *Journal of Health Policy Politics and Law* 33(2) 225-247.

REFERENCES

Culyer A. (2009) *Deliberative processes in decisions about health care technologies: combining different types of evidence, values, algorithms and people*. OHE Briefing no 48. London: Office of Health Economics.

Department for Communities and Local Government. (2009) *Multi-criteria analysis: a manual*. London: Department for Communities and local government. www.communities.gov.uk [Accessed 20th May 2010].

Department of Health. (2010a) *Equity and excellence: liberating the NHS*. Cm 7881. London: The Stationery Office. July 2010.

Department of Health (2010b) *A new value based approach to the pricing of new medicines*. A consultation. London: The Stationery Office. December 2010. www.dh.gov.uk/en/Consultations/Liveconsultations/DH_122760

Devlin N. (2010) The economics of a 'liberated' NHS. *Pharmacoeconomics*. 28 (12): 1075-1078.

Devlin N, Dakin H, Rice N, Parkin D, O'Neill P. (2011) *The influence of cost effectiveness and other factors on NICE decisions*. Paper presented at the Health Economics Study Group (HESG) conference, University of York, January 2011.

Dolan P, Shaw R, Tsuchiya A, Williams A. (2004) *QALY maximisation and people's preferences: a methodological review of the literature*. *Health Economics* 14 (2): 197-208.

Doyal I. (1997) The rationing debate: Rationing within the NHS should be explicit: The case for. *BMJ* 314:1114.

Gilovich T, Griffin D, Kahneman D. (2002) *Heuristics and biases: the psychology of intuitive judgement*. Cambridge University Press.

Goetghebeur M, Wager M, Khoury H, Rindress D, Gergoire J-P, Deal C. (2010) Combining MCDA, ethics and HTA: applying the EVIDEM decision-making framework to growth hormone for Turner syndrome patients. *Cost effectiveness and resource allocation*. 8:4. www.resource-allocation.com/content/8/1/4 [Accessed 12 August 2010].

Golan O, Hansen P, Kaplan G, Tal O. (2010) Health technology prioritization: which criteria for prioritizing new technologies, and what are their relative weights? *Health Policy* (forthcoming).

Golan O, Hansen P. (2008) *A new framework to help Israel's health basket committee choose new health technologies*. The VfM Chart. Report to the Israeli Health Basket Committee.

Hansen P, Ombler F. (2009) A new method for scoring multi-attribute value models using pairwise rankings of alternatives *Journal of Multi-Criteria Decision Analysis* 15, 87-107 <http://onlinelibrary.wiley.com/doi/10.1002/mcda.428/abstract> [Accessed 10 August 2010]

REFERENCES

- The Health Foundation. (2010) *Improvement in practice: commissioning with the community*. London: The Health Foundation, January 2010.
- HM Treasury. (2003) *The Green Book – Appraisal and evaluation in central government*. London: Her Majesty's Treasury.
- Highways Agency. (2010) *The appraisal process and Appraisal Summary Tables*. Highways Agency, 2010. www.highways.gov.uk/roads/15221.aspx [Accessed 2 July 2010].
- Highways Agency. (2009) *A46 Newark to Widmerpool improvement*. Highways Agency, June 2009. [www.highways.gov.uk/roads/documents/AST_12_June_2009_\(40.9KB\).pdf](http://www.highways.gov.uk/roads/documents/AST_12_June_2009_(40.9KB).pdf)
- Isle of Wight Primary Care Trust. (2009) *Commissioning strategy 2008-2013*. Version 2. Isle of Wight PCT, January 2009. <http://www.iow.nhs.uk/uploads/Commissioning/pdfs/CommissioningStrategy2008-2013.pdf>
- Kahneman D (2003) Maps of bounded rationality: psychology for behavioural economics. *American Economic Review* 93(5): 1449-1475.
- Kennedy I. (2009) *Appraising the value of innovation and other benefits: a short study for NICE*. London: NICE. www.nice.org.uk/media/98F/5C/KennedyStudyFinalReport.pdf [Accessed 6 August 2010].
- Leys M. (2003) Health care policy: qualitative evidence and HTA. *Health Policy* 65:217-226.
- Linkov I, Satterstrom F, Kiker G, Batchelor C, Bridges T, Ferguson E. (2006) From comparative risk assessment to MCDA and adaptive management: recent developments and applications. *Environmental International* 32: 1072-1093.
- Littlejohns P, Rawlins M. (2009) *Patients, the public and priorities in health care*. Oxford: Radcliffe.
- Martin D, Pater J, Singer P. (2001) Priority setting decisions for new cancer drugs: a qualitative case study. *The Lancet* 258:1676-81.
- Masya L, Young J, Solomon M, Harrison J, Dennis R, Salkeld G. (2008) An interactive computer based decision aid (Annalisa) balancing evidence and outcomes preferences to determine treatment options in rectal cancer. Poster presented at the COSA conference, 2008. http://cafeannalisa.org.uk/images/stories/annalisa/Lindy%20Annalisa_Poster_COSA_FINAL.pdf [Accessed August 13 2010].
- Mitton C, Smith N, Peacock S, Evoy B, Abelson J. (2009) Public participation in health care priority setting: a scoping review. *Health Policy* 91: 219-228.
- Morris S, Devlin N, Parkin D. (2007) *Economic Analysis in Health Care*. Wiley.

REFERENCES

- Mullen PM. Quantifying priorities in healthcare: transparency or illusion? *Health Services Management Research* 2004;17:47-58.
- Mussen F, Salek S, Walker S. (2007) A quantitative approach to benefit-risk assessment of medicines – part 1: The development of a new model using MCDA. *Pharmacoepidemiology and drug safety* 16: S2S15.
- Naden R, Barber A. (2010) *Ethical principles underlying explicit prioritization of patients for elective surgery*. Insight in: Devlin N, Appleby J. Getting the most out of PROMs: putting health outcomes at the heart of NHS decision-making. London: King's Fund & Office of Health Economics.
- National Commissioning Group. (2010) *Strengthening national commissioning: introducing the new decision-making framework for making recommendations for national commissioning*. www.ncg.nhs.uk
- New B. (1996) The rationing agenda. *BMJ* 312: 1593-1601.
- NICE. (2008a) *Guide to the Methods of Technology Appraisals*. Publication reference N1618. London: NICE.
- NICE. (2008b) Social value judgements: principles for the development of NICE Guidance. www.nice.org.uk/aboutnice/howwework/socialvaluejudgements/socialvaluejudgements.jsp [Accessed 6 August 2010].
- NICE. (2009a) Appraising life extending, end of life treatments. www.nice.org.uk/media/E4A/79/SupplementaryAdviceTACEoL.pdf [Accessed August 9th 2010].
- NICE. (2009b) *Guide to the single technology process*. Publication reference N201. London: NICE.
- Niezen MG, de Bont A, Busschbach J, Cohen JP, Stolk EA. (2009) Finding legitimacy for the role of budget impact in drug reimbursement decisions. *Int J Technol Assess Health Care* 25(1):49-55.
- Nobre F, Trotta L, Gomes L. (1999) Multi criteria decision-making – an approach to setting priorities in health care. *Statistics in Medicine* 18: 3345-3354.
- Peacock S, Mitton C, Bate A, McCoy B, Donaldson C. (2009) Overcoming barriers to priority setting using interdisciplinary methods. *Health Policy* 92: 124-132.
- Phillips, Lawrence D. (2006) *Decision conferencing*. Chapter 19 in: Operational Research working papers, LSEOR 06.85. Operational Research Group, Department of Management, London School of Economics and Political Science, London, UK. <http://eprints.lse.ac.uk/22712/> [Accessed August 10 2010]

REFERENCES

- Praet C, Shah K, Devlin N, Sussex J, Parkin D, Appleby J. (2011) *Is the aim of the health care system to maximise QALYs? An investigation of 'what else matters' in the NHS*. Paper presented at the Health Economics Study Group (HESG) conference, University of York, January 2011.
- Puig-Peiro R. (2011) *Economic appraisal across Government*. OHE Research Paper (forthcoming). (forthcoming).
- Rawlins M, Barnett D, Stevens A. (2010) Pharmacoeconomics: NICE's approach to decision-making. *British Journal of Clinical Pharmacology* 70(3):346-349.
- Royal National Orthopaedic Hospital NHS Trust. *Redevelopment of the RNOH – Outline Business Case*. Royal National Orthopaedic Hospital NHS Trust, May 2010.
http://rnoh.nhs.uk/sites/default/files/RNOH__OBC_-_Final_May_2010.pdf
- Ryan M, Gerard K, Amaya-Amaya M. (eds.) (2008) *Using discrete choice experiments to value health and health care*. Dordrecht: Springer.
- Scott SN, Lees A. (2001) Developing a prioritisation framework: experiences from a Scottish Health Authority. *Health Expectations* 4(1):10-17.
- Shah K. (2009) Severity of illness and priority setting in health care: a review of the literature. *Health Policy* 93:77-84.
- Stigler G. (1971) The theory of economic regulation. *Bell Journal of Economics and Management Science* 3: 3-18.
- Tappenden P, Brazier J, Ratcliffe J, Chilcott J. (2007) A stated preference binary choice experiment to explore NICE decision-making. 25(8):685-693.
- UK Border Agency (2010) *Tier 2 of the Points Based System – Policy guidance*. Tier 2 version 04/10. London: Home Office, UK Border Agency, 2010.
<http://www.ukba.homeoffice.gov.uk/sitecontent/applicationforms/pbs/tier2-guidance.pdf>
- Wilson E, Sussex J, Macleod C, Fordham R. (2007) Prioritizing health technologies in a Primary Care Trust. *Journal of Health Services Research and Policy* 12(2):80-85.

REFERENCES

Acknowledgements

The authors are grateful to Paul Hansen, Associate Professor, University of Otago (and developer of 1000Minds); Professor Donna Rindress (of The EVIDEM Collaboration) and Professor Jack Dowie (developer of Café Annalisa) for providing access to, and permitting reproduction of, the materials relating to their respective MCDA tools. Helpful comments on an earlier draft were provided by Tony Culyer, Jennifer Dixon and Martin Chalkley. Opinions expressed in this report are those of the authors.

