

Available online at www.sciencedirect.com

SciVerse ScienceDirect

journal homepage: www.elsevier.com/locate/jval

Multiple Criteria Decision Analysis for Health Technology Assessment

Praveen Thokala, PhD^{1,*}, Alejandra Duenas, PhD, MSc²

¹Health Economics and Decision Science, School of Health and Related Research, University of Sheffield, Sheffield, South Yorkshire, UK; ²IESEG, School of Management (LEM-CNRS), Paris, France

ABSTRACT

Objectives: Multicriteria decision analysis (MCDA) has been suggested by some researchers as a method to capture the benefits beyond quality adjusted life-years in a transparent and consistent manner. The objectives of this article were to analyze the possible application of MCDA approaches in health technology assessment and to describe their relative advantages and disadvantages. **Methods:** This article begins with an introduction to the most common types of MCDA models and a critical review of state-of-the-art methods for incorporating multiple criteria in health technology assessment. An overview of MCDA is provided and is compared against the current UK National Institute for Health and Clinical Excellence health technology appraisal process. A generic MCDA modeling approach is described, and the different MCDA modeling approaches are applied to a hypothetical

case study. **Results:** A comparison of the different MCDA approaches is provided, and the generic issues that need consideration before the application of MCDA in health technology assessment are examined. **Conclusions:** There are general practical issues that might arise from using an MCDA approach, and it is suggested that appropriate care be taken to ensure the success of MCDA techniques in the appraisal process.

Keywords: decision making, health economics, health technology assessment, multiple criteria decision analysis.

Copyright © 2012, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

Introduction

In the United Kingdom, the National Institute for Health and Clinical Excellence (NICE) makes recommendations to the National Health Service after assessing new and existing medical technologies. The current practice of NICE health technology appraisals is based on the incremental cost-effectiveness ratio (ICER), that is, the incremental cost per quality-adjusted life-year (QALY) gained by recipients of treatment. Even though NICE considers other criteria (e.g., severity and life saving) along with ICERs, there is concern that this approach may fail to capture other important sources of value [1–3]. In recognition of this issue, NICE commissioned Professor Sir Ian Kennedy to carry out a study on the relationship between innovation and the value of the technologies [4]. Also, recent developments such as the Patient Protection and Affordable Care Act in America [5] and the UK Department of Health's decision to use value-based pricing [6] indicate a paradigm shift toward transparency in using other criteria along with the traditional cost-effectiveness (C/E) analysis. Multicriteria decision analysis (MCDA) methods can support decision makers faced with evaluating alternatives by taking into account multiple criteria in an explicit manner [7]. They provide a structured and transparent approach to identify a preferred alternative by clear consideration of the relative importance of the different criteria and the performance of the alternatives on the criteria. In fact, a number of pharmaceutical drug manufacturers recommended the

use of MCDA (in their submissions to Professor Sir Ian Kennedy) but recognized that further research is needed before their implementation in the health technology appraisal process.

The main aspects of any MCDA method are 1) the alternatives to be appraised, 2) the criteria (or attributes) against which the alternatives are appraised, 3) scores that reflect the value of an alternative's expected performance on the criteria, and 4) criteria weights that measure the relative importance of each criterion as compared with others. MCDA approaches can be classified broadly into three categories: value measurement models, outranking models, and goal, aspiration, or reference-level models [7]. Figure 1 shows these three methods.

Value measurement models

The degree to which one decision option is preferred over another is represented by constructing and comparing numerical scores (overall value). The scores are developed for each individual criterion initially and aggregated into higher-level value models. Almost everyone who has suggested using MCDA methodology for health technology assessment (HTA) suggested this approach [8–11]; however, this approach is not without its constraints. Program budgeting and marginal analysis [12–14] and analytic hierarchy process [15,16], another widely used MCDA technique, are also similar to this value measurement modeling approach.

* Address correspondence to: Praveen Thokala, The University of Sheffield, HEDS, SchHARR, Regent Court, 30 Regent Street, Sheffield S1 4DA, UK.

E-mail: p.thokala@sheffield.ac.uk.

1098-3015 – see front matter Copyright © 2012, International Society for Pharmacoeconomics and Outcomes Research (ISPOR).

Published by Elsevier Inc.

<http://dx.doi.org/10.1016/j.jval.2012.06.015>

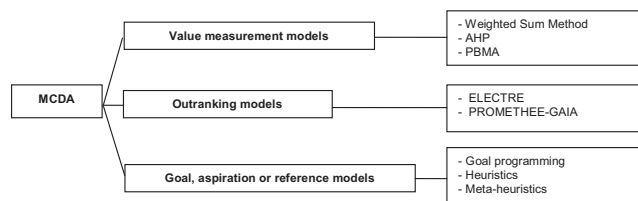


Fig. 1 – Classification of MCDA methods. AHP, analytic hierarchy process; ELECTRE, ELimination Et Choix Traduisant la REALité (ELimination and Choice Expressing Reality); MCDA, multicriteria decision analysis; PBMA, programme budgeting and marginal analysis; PROMETHEE-GAIA, preference ranking organization method for enrichment evaluations.

Outranking models

The alternatives are compared pairwise, initially in terms of each criterion, to assert the extent of preference for one over the other for that criterion. The preference information across all criteria is aggregated to establish the strength of evidence favoring the selection of one alternative over other. This approach is not widely used in health care but could also be an appropriate alternative for MCDA in HTA as it is based on direct comparison of the key characteristics of the alternatives.

Goal, aspiration, or reference-level models

This approach involves derivation of the alternative(s) that is closest to achieving the predefined desirable (or satisfactory) levels of achievement for each criterion [17]. Value-based pricing [18,19], a method to set the prices of drugs/treatments such that the ICER is under the relevant C/E threshold, can be based on mathematical programming techniques similar to those used in this MCDA approach.

Incorporating multiple criteria in health technology appraisals

Health policy decision makers internationally so far have been considering cost-per-QALY ratios alongside other criteria, such as equity and fairness and prioritization of interventions for vulnerable populations, in a deliberative manner [20]. An integrated ICER that includes other sources of value has been proposed to allow explicit incorporation of other criteria, such as societal preferences, disease severity, equity and benefits to caregivers, in the existing ICER framework. Societal preferences relating to distributional justice captured from surveys have been proposed for inclusion in the ICER calculations [21,22]. Explicit incorporation of equity in calculating ICERs for HTAs has also been considered [2,23,24], but a need for further research has been identified [25]. A hybrid method that supplements the current ICER evaluation for NICE with a comprehensive benefits and value review has also been proposed [1,8]. This approach attempts to capture the sources of value not systematically considered at the present (such as innovation, societal benefit, disease severity, unmet need, patient compliance, and related benefits) by using different ICER thresholds for different comprehensive benefits and value scores [8].

Health care organizations in a few countries have made attempts to incorporate different criteria into their decision-making processes. For example, the Netherlands health care system used four criteria (care must be necessary, effective, and efficient and cannot be left to the individual's own responsibility) for priority setting, but it made the decision-making process rather complex [26]. Some countries such as Belgium and France, in an attempt to establish a relationship between financial constraints and medical need, have proposed to vary their pharmaceutical expenditure on the basis of patients' medical need [27,28]. France classifies med-

icines into three categories: essential, important, and convenience, and French people receive 100% reimbursement for essential drugs, 65% for important drugs, and 35% for convenience medication [28]. In South Korea, the Health Insurance Review and Assessment service considers clinical benefit, C/E, budget impact, reimbursement status in other countries, and other features that might affect public health in determining whether a new drug receives reimbursement [29].

MCDA has also been used to inform health care decisions [30–32], setting priorities for HTAs [33], and other governmental issues [34,35]. The benefit-risk assessment of medicines, based on multiple benefit and risk criteria including the trade-offs between the benefits and the risks, was also performed by using MCDA [36,37]. MCDA techniques have also been used for shared decision making between patients and doctors in the evaluation and selection of therapies, treatments, and health care technologies [38,39]. These MCDA techniques were said to identify and include the personal preferences of the patient, but the complexity of the MCDA models and the time taken to complete the model were mentioned as disadvantages [40,41]. Program budgeting and marginal analysis [13,14,42], used for reallocation of scarce health care resources, is similar to MCDA methodologies. This method has received some attention in the health sector [12,43], but its success has been limited because of the complexity of the approach, large data requirements, and organizational barriers [44,45].

Despite the widespread use of MCDA in other health streams, it is only recently that there have been studies that advocate the use of MCDA for HTA. A framework utilizing a value matrix was developed to include quantifiable components that are currently considered in health decision making to promote transparent and efficient health care decision making [9]. This framework was also linked to a qualitative assessment including six ethical and health system-related components of decision to provide a tool for combining HTA, MCDA, values, and ethics [10]. Health England Leading Prioritisation study also used MCDA to prioritize investment in preventative health interventions [46].

Most of the proposed MCDA approaches, however, use the same technique (weighted sum approach), which may lead to the researchers/health professionals assuming that it is the only relevant MCDA method. This article attempts to provide an overview of the main MCDA methods available and the issues with their implementation in a technology appraisal process.

MCDA versus NICE Appraisal Process

This section compares the MCDA approach to the NICE appraisal process. Although NICE is chosen as the example, the findings are generalizable to other international health care decision-making organizations. MCDA is aimed at supporting decision makers faced with evaluating alternatives taking into account multiple and often concurrent criteria. The MCDA process consists of the following phases: problem identification and structuring, model building and use, and the development of action plans [7]. The appraisal process followed by NICE is divided into three phases: scoping, assessment, and decision making through deliberation by a committee that makes its recommendations on the basis of evidence and experts' and patients' opinions.

The MCDA process is compared with the current NICE technology appraisal process as shown in Figure 2. The current NICE approach includes this problem-structuring process during the "scoping" stage to set the predefined options (treatments, drugs, etc.) and the key outcomes relevant for the appraisal process. The criteria for NICE appraisals are defined in the methods guide [47], not separately for each appraisal, but the scoping process allows identification of other key issues (such as disease-specific outcomes). The first two steps of the MCDA process, that is, identifying alternatives and criteria, is known as problem structuring; this

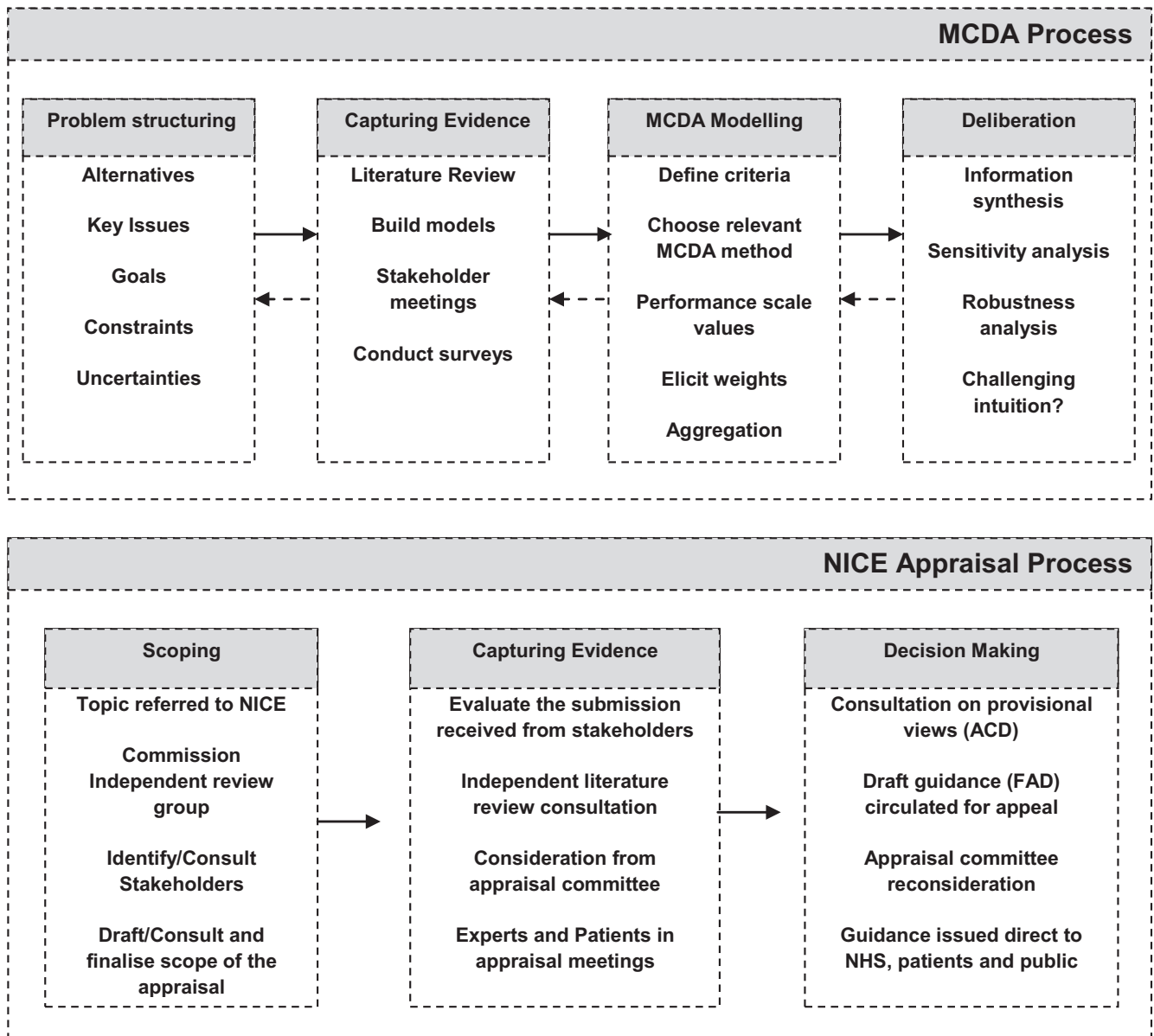


Fig. 2 – MCDA and NICE technology appraisal processes. ACD, appraisal consultation document; FAD, final appraisal determination; MCDA, multicriteria decision analysis; NHS, National Health Service; NICE, National Institute for Health and Clinical Excellence.

is usually achieved by decision conferencing [48,49], which involves the meeting of all the relevant stakeholders. It is anticipated, however, that the criteria included in the MCDA approach would be based on criteria used by NICE. Thus, there is not much difference between the current NICE scoping approach and the first two steps of the MCDA process. It should be noted that the current criteria used by NICE or other organizations may not fulfill all the requirements of MCDA methods such as nonredundancy, judgmental independence, completeness, operationality, and measurability [7]. The criteria should be analyzed to ensure that they are suitable for use in MCDA modeling, for example, ensuring that the criteria are not redundant to avoid double counting (e.g., efficacy/effectiveness and cost/effectiveness) when using value measurement models.

It is in the decision-making stage that the MCDA and the current NICE appraisal processes differ [50]. In the NICE approach, the evidence regarding the alternatives is captured and

presented as a report (including a table of summary characteristics comparing the interventions on the key criteria). An appraisal committee then makes a decision in a deliberative manner by using ICER and other criteria. In the MCDA approach, this evidence needs to be quantified and input into mathematical models to identify the best alternative(s). The manner in which these models are built separates the different MCDA techniques. Thus, criteria definition process is the primary starting point for both approaches, and an MCDA model would build on the existing criteria that NICE is using by applying the design principles of MCDA. MCDA approaches can support rather than replace the deliberative process already existent in NICE by adding consistency and transparency through explicit scoring and weighting of criteria. It should be noted that another key difference between the current NICE technology appraisal process and the proposed MCDA approaches is the principle of “opportunity cost.” NICE is charged with value for money across all of

Table 1 – Characteristics of the drugs in the appraisal process.

Characteristics	Drug A $z_i(a)^*$	Drug B $z_i(b)^*$
C/E (in terms of NB)	£15,850	£25,600
Equity (%)	0.14	0.08
Innovation	Innovative	Less Innovative
Patient compliance (%)	0.93	0.85
Quality of evidence	Good	Good

C/E, cost-effectiveness; NB, net benefit.
^{*} $z_i(a)$ and $z_i(b)$ are attribute values on the *i*th criterion for drug A and drug B, respectively.

health and social care; it needs to consider the costs and benefits foregone from displaced activities and does so by using the ICER threshold approach. The current ICER threshold approach is based on using QALY as the measure of benefit and if other criteria are to be included in benefits, the value of threshold will need to be recalculated, that is, the opportunity costs of the displaced intervention will be different with an MCDA approach than with a cost-per-QALY approach as the opportunity cost needs to be considered across all criteria for displaced activities. Furthermore, this approach to estimate C/E, assessing the benefits of a technology against the ICER threshold, could be interpreted as a decision-making framework than as a criterion (or an aspect of benefit) in an MCDA model. Thus, this article does not include opportunity costs and the case study is set up as a decision between two drugs using the net benefit (NB) approach.

(A Hypothetical) Case Study

One of the objectives of this article was to analyze the possible application of MCDA methods in HTAs. A case study is presented in this section to show how different MCDA techniques can be applied to NICE technology appraisals.

The case study presented in this section is based on a hypothetical NICE technology appraisal process where a recommendation is needed to be made between two drugs A and B, where drug A is the current intervention and drug B is the new intervention. The characteristics of each of the drugs when compared against the best standard care are shown in Table 1. The C/E of the drugs is measured by using NB calculated assuming a willingness-to-pay threshold of £20,000 per QALY; if C/E is used to make the decision, drug B would be recommended over drug A; however, MCDA could also be used to compare the drugs.

The three MCDA approaches mentioned in the Introduction section have been applied to this case study and are described below. Although the case study uses two interventions (current intervention and new intervention), all the MCDA approaches can be extended to include more interventions. It should be noted that the criteria specified here are for demonstration purposes only. The aim of this study was not to enter into a debate on what criteria should be used and their definitions but to compare different MCDA techniques for incorporating multiple criteria into the decision process once the relevant criteria are identified.

Before the MCDA models can be developed, the performance of the alternatives (drugs A and B) against the specified criteria needs to be measured in an objective manner. To achieve this, measures (or scales) that can describe the desirability of achieving different levels of performance for each criterion need to be identified. For some criteria, such as patient compliance, where preferences might be assumed to be linearly related to the attribute's value, the attribute value z_i can be substituted for the performance on the

criterion. In most cases, however, this needs to be modeled as there is rarely such a simple linear relationship between attribute values (z_i) and preferences. In such cases, a scale needs to be constructed to represent the performance of alternatives; it should be noted that choosing the scales (usually ordinal or ordered-categorical scales) to model these performance measures is not trivial. In this article, it is assumed that scales to measure the performance of the drugs on various criteria already exist and are specified. Furthermore, it is assumed without loss of generality that all these scores are defined in such a manner that increasing values are preferred.

Henceforth, the performance levels of drugs A and B on the *i*th criterion as measured on these scales are represented as performance score values $v_i(a)$ and $v_i(b)$, respectively. For any criterion *i*, performance score $v_i(a)$ is a nondecreasing function of the attribute value $z_i(a)$; this could also be defined more generally for any criterion *i* as $v_i = f(z_i)$ as the function *f* is the same for all alternatives (drugs) keeping in line with the need for the performance of the different alternatives to be measured in an objective manner. The performance value scores are generally standardized to 0 at worst outcome and to a convenient value (usually 1, 10, or 100) at the best outcome.

The three different MCDA approaches mentioned are applied to this case study as below; this allows us to demonstrate the potential advantages and pitfalls of using the different MCDA modeling approaches.

Value measurement models

This approach is based on constructing a single overall value for each alternative to establish a preference order of alternatives. An alternative A is said to be preferred to B if $V(a) > V(b)$, where $V(a)$ and $V(b)$ are overall values (taking into account all *n* criteria) of A and B, respectively. Also, there is said to be indifference between the alternatives if $V(a) = V(b)$.

The first step in this approach is to do preference modeling, that is, constructing the performance levels of drugs A and B on all criteria as shown in Table 2. The performance score values $v_i(a)$ and $v_i(b)$ of drugs A and B on the *i*th criterion, also known as partial value functions, are bounded between 0 (worst outcome) and 1 (best outcome). The importance of different criteria can be measured by using the gain associated with replacing the worst outcome by the best outcome, and the weights w_i represent the relative importance of the *i*th criterion. Weights are assigned independently of the alternatives to provide consistency across comparisons, and scores for each criterion are then assigned for each alternative. Furthermore, the weights can be normalized to allow interpreting of the weight of individual criterion as a proportion of the total weight. The final step is to aggregate these partial value functions taking into account the relative importance of dif-

Table 2 – Performance levels of drugs.

Criterion (i)	Drug A $v_i(a)^*$	Drug B $v_i(b)^*$	Weights w_i
C/E	0.72	0.84	8
Equity (%)	0.14	0.08	1
Innovation	0.91	0.62	3
Patient compliance (%)	0.93	0.85	2
Quality of evidence	0.82	0.79	3

C/E, cost-effectiveness.

^{*} $v_i(a)$ and $v_i(b)$ are performance value scores on criterion *i* for drug A and drug B, respectively.

[†] w_i is the weight of criterion *i*.

ferent criteria; the manner in which this is done separates different value measurement approaches.

In this article, additive aggregation (also known as weighted sum approach) is described as it is the most common value measurement modeling approach [7] and it is based on the following equation:

$$V(a) = \sum_{i=1}^n w_i v_i(a) \tag{1}$$

where $V(a)$ represents the overall value, w_i represents the relative importance, and $v_i(a)$ represents the score of alternative A on the i th criterion, respectively.

This approach requires some assumptions regarding the criteria and their weights, namely, preferential independence of criteria and the need for the weights to satisfy the trade-off requirements. Preferential independence requires that the decision can be made by using a subset of criteria if the other criteria are the same for all alternatives irrespective of their actual values; that is, the decision can be made by using only the criteria on which the alternatives differ. The weight parameters w_i also need to follow a strict trade-off condition to capture the concept of “importance” as well as compensate for the different measurement scales of different criteria. This is achieved by swing weights, which represent the gain in overall value by going from the worst value to the best value in each criterion; that is, for any two criteria i and k , the ratio w_i/w_k is the change in $v_k(a)$ that should compensate for a unit loss on $v_i(a)$. There are a number of ways in which these swing weights can be elicited; these techniques are not discussed here as they are explained in detail in the literature [7].

In this case study, it is assumed that the relative weights and the performance of the alternatives on different criteria have been identified by using appropriate techniques and are as shown in Table 2. Using this information, the overall values of drugs A and B can be calculated as follows:

$$V(a) = \sum 8 \times 0.72 + 1 \times 0.14 + 3 \times 0.91 + 2 \times 0.93 + 3 \times 0.82 = 12.95 \tag{2}$$

$$V(b) = \sum 8 \times 0.84 + 1 \times 0.08 + 3 \times 0.62 + 2 \times 0.85 + 3 \times 0.79 = 12.73 \tag{3}$$

If the decision was made by using this approach, drug A would be recommended over drug B as the aim of this approach is to identify the alternative with maximum value. Drug A has a higher overall value due to the higher weights w_i placed on innovation, compliance, and quality of evidence in which drug A performs better than drug B.

This approach is simple to use, but as observed in this scenario, poor performance on a criteria (C/E) can be overcome by doing well in other criteria depending on the weights and partial value functions. Also, considerable caution needs to be taken to satisfy the

Table 4 – Outranking relations and weights.

Criterion (i)	Weights ω_i^*	Drug A	Drug B
C/E	10		✓
Equity (%)	2	✓	
Innovation	1	✓	
Patient compliance (%)	3	✓	
Quality of evidence	2	—	—

* ω_i is the weight of criterion i .

preferential independence of criteria and the corresponding trade-offs of swing weights.

Outranking approach

This principle of outranking is based on the general concept of dominance [51,52]. If the performance of two alternative drugs A and B on each of the i criteria is $v_i(a)$ and $v_i(b)$, respectively, we can conclude that drug A should be preferred over drug B if $v_i(a) \geq v_i(b)$ for all criteria (with strict inequality for at least one criterion). In this event, drug A is said to dominate drug B. Strict domination, however, rarely occurs in practice, and thus the evidence needs to be evaluated in a systematic manner. More generally, drug A outranks alternative drug B if there is sufficient evidence to justify a conclusion that drug A is at least as good as drug B, taking all criteria into account.

This approach utilizes outranking relation (i.e., comparing performance scores on individual criterion to see which alternative outranks the other on that criterion) on a set of alternatives focusing on pairwise comparisons, and these pairwise comparisons are used to estimate the concordance and discordance indices. For drug A, concordance index is the evidence in favor of A outranking B while the discordance index is the evidence against A outranking B. Similarly, for drug B, the concordance index is the evidence in favor of B outranking A while the discordance index is the evidence against B outranking A.

The first step in estimating the concordance and discordance indices is to construct a matrix of outranking relations from the individual scores on each criterion. The performance scores of drugs against the individual criteria are shown in Table 3 while the matrix of outranking relations along with the relative weights for different criteria is shown in Table 4. The outranking approach recognizes that performance scores, $v_i(a)$ and $v_i(b)$, are imprecise measures, and so alternative a is preferred to alternative b only if $v_i(a) - v_i(b)$ exceeds a predefined “indifference threshold.” For example, if the threshold was 0.05, alternative drug A and drug B would be incomparable on “quality of evidence” criteria as the difference 0.03 is less than the threshold. The indifference thresholds can be specified for the difference in either performance value scores $v_i(a)$ or the attribute values $z_i(a)$. Also, it is to be noted that the weights of different criteria are uninfluenced by the scale of the value functions and thus do not need to follow the theoretical concept of trade-offs as required by the value measurement approach; they just represent the relative importance of different criteria in the assertion that one alternative is better than the other.

There are a number of ways to quantify concordance and discordance indices that correspond to different outranking methods. In this study, ELECTRE I [53] is used but the reader should bear in mind that there are a number of other options (ELECTRE II, III, IV, TRI [51,52,54], PROMETHEE [55], and GAIA [56]). In ELECTRE I, the concordance index is defined as the ratio of the sum of weights in the criteria for which drug A is at least as good as drug B to the sum of weights in all criteria, that is,

$$C(a, b) = \frac{\sum_{i \in Q(a, b)} \omega_i}{\sum_{i=1}^m \omega_i}$$

Table 3 – Performance scores of drugs.

Criterion (i)	Drug A $v_i(a)^*$	Drug B $v_i(b)^*$
C/E	0.72	0.84
Equity (%)	0.14	0.08
Innovation	0.91	0.62
Patient compliance (%)	0.93	0.85
Quality of evidence	0.82	0.79

C/E, cost-effectiveness.
* $v_i(a)$ and $v_i(b)$ are performance value scores on criterion i for drug A and drug B, respectively.

Table 5 – Attributes of drugs, the goals and weights against different criteria.

Criterion (i)	Drug A $z_i(a)^*$	Drug B $z_i(b)^*$	Goals g_i	Weights $\omega_i^{+\dagger}$	Weights $\omega_i^{-\dagger}$
C/E	£15,850	£25,600	£20,000	10	0
Equity (%)	0.14	0.08	0.20	0	5
Innovation	Innovative	Less Innovative	—	0	0
Patient compliance (%)	0.93	0.85	0.95	0	5
Quality of evidence	Good	Good	—	0	0

C/E, cost-effectiveness.
 * $z_i(a)$ and $z_i(b)$ are attribute values on criterion i for drug A and drug B, respectively.
 † Ω^{+i} and Ω^{-i} are the weights attached to the deviations from the goal g_i for criterion i .

where $Q(a, b)$ is the set of criteria for which A is at least as good as B. For the discordance index, a veto threshold t_i can be specified as follows:

$$D(a, b) = \begin{cases} 1 & \text{if } v_i(b) - v_i(a) > t_i \text{ for any } i \\ 0 & \text{otherwise} \end{cases}$$

The concordance and discordance indices are compared against the concordance (C') and discordance (D') thresholds, respectively, to estimate the outranking relation. If the concordance index (C) is greater than the concordance threshold (C') and the discordance index (D) is less than the discordance threshold (D'), then that drug is said to outrank the other drug. If the thresholds are specified such that both drugs outrank each other, then they are said to be indifferent. In case of neither drug outranking the other, the drugs are said to be incomparable. If there are more than two alternatives (i.e., A, B, C, etc.), the concordance and discordance indices are estimated for each pair to build an outranking relation by using the concordance and discordance thresholds C' and D' , respectively.

In the current case study, the concordance index of drug A against drug B is the sum of weights in $Q(a, b)$, set of criteria for which A is at least as good as B, divided by the overall sum of weights; that is, $(2 + 1 + 3 + 2)/(10 + 2 + 1 + 3 + 2) = 8/18$. The decision, however, can be vetoed by the poor performance of drug A in C/E by specifying the veto threshold t_1 as 0.1 for C/E (drug B is better than A in C/E by 0.12, which is greater than the veto threshold t_1). Similarly, the concordance index of drug B against drug A is 10/18. If the concordance threshold (C') is less than 0.56, drug B is said to outrank drug A provided the veto thresholds for other criteria are not violated as the concordance index is greater than the concordance threshold.

This method does not need the theoretical requirement of trade-offs for weights as required in the value measurement models; the weights just convey the relative importance of the different criteria. This method is intuitive, and the use of indifference and veto thresholds allows more flexible/realistic decision rules to be specified. There are different levels of complexity—ELECTRE I, II, III, IV, TRI, PROMETHEE, and GAIA—in the outranking approach. This approach might lead to incomparability if two drugs are quite similar; however, one could argue that this is appropriate for the appraisal process as further deliberation might be needed to choose between the drugs if their performance is quite similar.

Goal programming

Goal programming involves a mathematical formulation of the satisficing heuristic; the term “satisficing” is a combination of the terms “satisfy” and “suffice.” The emphasis of the satisficing model is on attaining satisfactory levels of performance on each criterion, considering the preference of criteria in their order of importance. Satisficing levels are predefined as “goals,” and a programming algorithm is used to identify the alternatives that satisfy the goals in the specified priority order [57].

Unlike the weighted sum approach, which involves developing partial value functions $v_i(a)$, the goal programming method operates directly on the attribute values, $z_i(a)$, of the alternatives on the criteria, as it is more operationally meaningful to match measurable attributes to the goals. The attribute values of alternative A corresponding to the “ n ” criteria are represented as $z_1(a), z_2(a), \dots, z_n(a)$ while the “goals” for each criterion are represented as g_1, g_2, \dots, g_n as shown in Table 5. These goals (“aspiration levels”) are usually defined by the decision maker and are understood as desirable levels of performance for each attribute value.

Direction of preference in a goal programming context represents the relationship between the attribute value and the goal. Three alternatives for the direction of preferences can be found: 1) when the attribute is maximized, the goal’s level of performance achieves a minimum representing a point of “satisfaction,” for example, attaining at least 95% patient compliance; 2) when the attribute is minimized, the goal’s level of performance achieves a maximum representing a point of “satisfaction,” for example, the ICER threshold of no more than £20,000 per QALY; and 3) when the decision makers define a most desirable level of performance for the attribute that must be as close to the goal as possible. The difference between the attribute values and the goals are represented as goal deviations d_i^+ or d_i^- , that is, the amounts a targeted goal is exceeded or underachieved, respectively.

Goal programming involves minimizing the goal deviations taking the relative importance of goals into account. There are two main variants of goal programming [58]—weighted goal programming and lexicographic goal programming; they differ in the way the optimal solution is prioritized and achieved. The weighted goal programming approach minimizes the unwanted deviations after assigning weights to the goal deviations according to their relative importance as shown in the following equation:

$$\min D = \sum_{i=1}^n (\omega_i^+ d_i^+ + \omega_i^- d_i^-)$$

subject to $f_i(x) - d_i^+ + d_i^- = g_i$ for $i = 1, \dots, n$

where x is the set of decision variables (independent) to be determined, $f_i(x)$ is the attribute value $z_i(a)$ as a function of the independent variable(s), g_i is the target value for the i th criterion, d_i^+ and d_i^- are the negative and positive deviations from this target value, and ω_i^+ and ω_i^- are the respective weights attached to these deviations. The lexicographic goal programming formulation orders the goals into a number of priority levels and minimizes them in a lexicographic manner, that is, deviation in a higher priority level being more important than any deviations in lower priority levels. This sequential minimization approach minimizes each priority while maintaining the minimal values reached by all higher priority level minimizations by adding them as explicit constraints.

In this case study, it was assumed that patient compliance and equity are difficult to change but C/E can be improved by changing the price of the drug (akin to value-based pricing). As C/E is the only attribute that can be changed, it does not matter whether a

weighted goal programming (GP) approach or a lexicographic GP approach is utilized. In this study, the lexicographic GP approach is used with C/E as the highest priority and all the other criteria together as the next priority. Assuming the NB for drug A varies according to $f(x) = 25,000 - 1000x$ (where x is the unit price of drug A), the price of drug A has to be decreased by 45% (from the initial price of £9.15 to £5.00) so that the C/E goal of an NB of £20,000 is achieved. The NB for drug B is already above the specified goal threshold. In reality, an NB higher than £20,000 would be encouraged by the decision makers but for this case study the NB target represents the minimum requirement.

Now that both drugs have achieved the target C/E, the analysis can move to the next priority level, which includes all the other criteria. From Table 5, it can be seen that the only two criteria that have a weight value other than zero are equity ($w_2 = 5$) and patient compliance ($w_4 = 5$); therefore, the negative deviations d_2^- and d_4^- are to be calculated as follows:

Drug A

$$d_2^-(a) = g_2 - z_2(a) = 0.2 - 0.14 = 0.06 \tag{4}$$

$$d_4^-(a) = g_4 - z_4(a) = 0.95 - 0.93 = 0.02 \tag{5}$$

Drug B

$$d_2^-(b) = g_2 - z_2(b) = 0.2 - 0.08 = 0.12 \tag{6}$$

$$d_4^-(b) = g_4 - z_4(b) = 0.95 - 0.85 = 0.1 \tag{7}$$

Thus, the deviations for drug A and B are

$$D(a) = (5 \times 0.06 + 5 \times 0.02) = 0.4 \tag{8}$$

$$D(b) = (5 \times 0.12 + 5 \times 0.1) = 1.1 \tag{9}$$

Drug A performs better than drug B in terms of getting closer to the equity and compliance goals; thus, it could be recommended on the condition that its price is reduced by 45% (to ensure that drug A satisfies the C/E goal).

Another variation to this approach is to have a range of goals based on other criteria. For example, the C/E could have different thresholds based on the alternative's performance on other criteria [8]. In practice, this could be implemented by using a higher threshold if the aggregated goal deviations for other criteria are low; that is, a technology could be assigned a higher ICER threshold (as it performs better in achieving close to other goals).

This goal programming approach echoes similarities with value-based pricing [18,19], in which mathematical programming techniques can be used to estimate the price of the drug, based on the definition of "value" chosen by the health organizations. Also, obviously, $f(x)$ will never be as simple as our assumption, and so complex C/E models need to be built and analyzed to identify the price of the drug such that the ICER is under the recommended value-based threshold. It should be noted that a number of heuristics have been developed to deal with complexity as the computational time for a goal programming problem is directly related to complexity.

Comparing Different MCDA Approaches

Table 6 compares the different MCDA approaches on a number of dimensions to provide an indication of the potential benefits and limitations of each of the approaches. First, there are different requirements of the weights depending on the MCDA approach, with value measurement models requiring additional effort compared with outranking and goal programming methods because of the time needed to interpret swing weights. Similarly, value measurement models need significant effort to develop the performance value scores while the goal programming and outranking methods can be implemented on the attribute values directly. Value measurement models, however, are easy to understand and can enable real-time sensitivity analysis. Both outranking and goal programming methods are easy to follow, but significant computational time is needed for goal programming. Similarly,

Table 6 – Comparison of different MCDA approaches.

	Value measurement models	Outranking approach	Goal programming
Weights	Swing weights are used to capture both the effect of measurement scales and the importance of the criteria Weights need to satisfy the preferential independence of criteria and the trade-off requirements	Weights are uninfluenced by the scale of the value functions. They convey the relative importance of criteria in the assertion that one alternative is better than the other Weights do not have to satisfy any conditions	Weights are attached to the deviations and represent the relative importance of criteria by specifying an overall measure of deviations from the goals Weights do not have to satisfy any conditions
Measuring the performance of the criteria	Performance scores $v_i(a)$, monotonic functions of the attribute values $z_i(a)$, need to be developed for all criterion i . Significant effort is needed to develop these performance scores	Outranking approach can use either performance value scores $v_i(a)$ or the attribute values $z_i(a)$, saving on the effort needed to develop performance scores	Goal programming method operates directly on the attribute values, $z_i(a)$. No need to develop performance scores.
Complexity of the MCDA model	Weighted sum approach is easy to understand and use by the decision makers. The parameters can be changed in real time to observe their effect	Intuitive and easy to follow. With right software, assumptions can be changed and results can be observed almost instantaneously	Easy to understand but requires significant computational time to provide results. Real-time updating is not possible
Presentation of the results	Easy to follow and enables further deliberation, well suited for good visual presentation of the results	Moderately easy to follow, can be presented visually but difficult with multiple alternatives	Results easy to follow, but they cannot be represented visually
Incorporating uncertainty	Probabilistic sensitivity analysis can be used to propagate parameter uncertainty quite easily.	Moderately difficult to include uncertainty, needs specialist software	Quite difficult to include uncertainty, complex stochastic programming techniques are needed
MCDA, multicriteria decision analysis.			

results from value measurement models lend themselves for easy visual presentation while results from outranking and goal programming methods are difficult to follow. Finally, uncertainty is easier to incorporate in value measurement models than in outranking or goal programming approaches.

Generic Issues with Using MCDA

The potential issues that might arise with implementing any MCDA approach in the HTA process are detailed in the following subsections.

Appraisal-specific or generic process

The MCDA process could be the same for all technology appraisals or it could be tailored to a given appraisal under consideration. For example, the criteria (and their weights) can be different for each appraisal or the same set of criteria can be used for all appraisals. Furthermore, the functions that estimate the value of alternatives against the criteria could be fixed for all the appraisals or appraisal-specific functions could be built on the basis of appraisal characteristics. The chosen MCDA method needs to be transparent, consistent, auditable, and defensible as it is a national decision-making process. Thus, appropriate care needs to be given before deciding on an appraisal-specific process or a standard approach for all the appraisals. Committee members' preferences or societal preferences need to be used as weights on the basis of whether an appraisal-specific or generic MCDA process is chosen, respectively.

If an appraisal-specific MCDA approach is chosen, the weights are elicited from the committee members. The weights (i.e., preferences of the individual decision makers) can be captured in a workshop setting by using deliberation, by direct rating of alternatives (such as visual analogue scales analytic hierarchy process, ranking, and point allocation), or by indirect weight elicitation methods such as discrete choice experiments [59–61]. As the decision committee includes a number of individuals from different perspectives (patients, clinicians, administrators, etc), there can be issues with group dynamics in the collection and aggregation of the individuals' preferences (i.e., criteria weights). Variations are expected across decision makers according to their perspective and value systems (e.g., a patient might see safety as more important than efficacy and the contrary for a clinician). MCDA can address this variation by allowing each individual on a committee to express his or her perspective and aggregating the weights, with the method for aggregation of individuals' preferences dependent on whether a consensus needs to be achieved by the committee. If a consensus is necessary, the individuals in the committee need to share/compare their values to identify issues of conflict and achieve common ground, which can be difficult to achieve. Otherwise, the overall value can be calculated as an average of the individual values, which could be anonymous if need be. This variation can be represented as the SDs associated with the mean values of the weights, and sensitivity analysis can be performed to see the robustness of results to changes in the weights.

If a generic MCDA approach is chosen, the weights need to be elicited from the general population, which requires a large number of resources to capture the population preferences. The selection of the criteria needs to be made a priori by the policymakers. Furthermore, it needs to be ensured that the general population understands (and interprets) the meaning of the weights and the value scores correctly. Consistency in the MCDA process across different appraisals and over time can be achieved by clearly defined criteria and the weights estimated from the public in advance. In this approach, the role of the appraisal committees for a given appraisal would be to interpret the evidence regarding characteristics of the relevant diseases, patients, and interventions of

interest and to assess which of the criteria apply and how they should be valued. The decision makers use MCDA to evaluate the data and context to come up with a decision, which is similar to the deliberative process in current practice. Given the considerable resources needed to estimate general population preferences, it might be better to build the population consultation on some exploration at the committee level, provided the committees are a good representation of the general population.

Uncertainty modeling

There are three main areas of uncertainty involved with using MCDA in the HTA process, namely, uncertainty in problem structuring (i.e., choosing the right MCDA model, criteria, level of detail, etc.); uncertainty with evidence of different alternatives; and variation in preferences (i.e., uncertainty in performance scores, criteria weights, thresholds, etc.). Structural uncertainty is hard to capture and is out of the scope of this article. The uncertainty in clinical effectiveness and C/E as well as other evidence (usually caused by extrapolating the data from a randomized controlled trial to a general population) has a direct effect on committee members' preferences; thus, it should be ensured that they understand this uncertainty in evidence. Scenario analyses [62], multiattribute utility theory [63], fuzzy logic [64], and stochastic multicriteria acceptability analysis [65] can be used to capture this uncertainty. Variation in criteria values, weights, and thresholds evident during the aggregation of the preferences of individuals in the decision committee can be represented by using SDs associated with the mean values. Sensitivity analysis can be performed to see the robustness of results to changes in the model parameters. Because of the interdependence of uncertainty with evidence of different alternatives and variation in committee members' preferences, appropriate care needs to be taken in performing uncertainty analysis. Probabilistic sensitivity analysis can also be used to capture and propagate the parameter uncertainty with the help of Monte-Carlo simulation techniques [66,67].

Practical issues

The practicality of applying MCDA will involve a significant workload/cognitive burden on decision makers. Appropriate thought needs to be given to the practical aspects such as whether to train all the committee members in the relevant techniques of MCDA or whether to have a facilitator(s) to help use the techniques in the decision process. Also, the MCDA techniques rely on preference capturing, statistical analysis, and synthesizing data, which may require specialist software or programs; thus, the relevant software/program requirements need to be identified. This also relates to other practical issues such as the methods of data capturing (survey sheets on paper, computer-based forms, etc.) and data aggregation. Data aggregation involves capturing the individual committee members' preferences and transferring them into appropriate software; this could be done in real time or in between the meetings. The MCDA model developed needs to be explored to ensure the robustness of key factors; this can be performed premeeting, during the meeting, or postmeeting depending on the availability of the MCDA facilitator. Finally, the model outputs need to be visualized and incorporated into the final documentation along with the recommendations. All these aspects involve significant burden on decision makers, and it needs to be decided whether the transparency and consistency achieved by using the MCDA process is worth the additional burden.

Conclusions

An overview of the MCDA process is provided and is identified to be similar to the existing NICE appraisal process but with the addition of a formal mathematical approach to decision making. The main MCDA modeling approaches are applied to a hypothetical case study, and their potential strengths and weaknesses are outlined. The potential users need to understand the general practical issues that might arise from using an MCDA approach in the HTA process and choose an appropriate MCDA method to ensure the success of MCDA techniques in the appraisal process.

Source of financial support: This research has received financial support from the NICE Decision Support Unit, but NICE had no responsibility or control over the content of this article.

REFERENCES

- [1] Goldman D, Lakdawalla D, Philipson TJ, Yin W. Valuing health technologies at NICE: recommendations for improved incorporation of treatment value in HTA. *Health Econ* 2010;19:1109–16.
- [2] Baeten SA, Baltussen RM, Uyl-de Groot CA, et al. Incorporating equity-efficiency interactions in cost-effectiveness analysis—three approaches applied to breast cancer control. *Value Health* 2010;13:573–9.
- [3] Klein R. Rationing in the fiscal ice age. *Health Econ Policy Law* 2010;5:389–96.
- [4] Professor Sir Ian Kennedy. *Appraising the Value of Innovation and Other Benefits: A Short Study for NICE*. United Kingdom: National Institute for Clinical Excellence, 2010.
- [5] Neumann PJ, Weinstein MC. Legislating against use of cost-effectiveness information. *N Engl J Med* 2010;363:1495–7.
- [6] Department of Health. *A New Value-Based Approach to the Pricing of Branded Medicines—A Consultation*. 2010.
- [7] Belton V, Stewart TJ. *Multiple Criteria Decision Analysis: An Integrated Approach*. Kluwer Academic Publishers, 2002.
- [8] Precision Health Economics LLC. *Submission to the Kennedy Review*. Santa Monica, California: Precision Health Economics, 2009.
- [9] Goetghebeur MM, Wagner M, Khoury H, et al. Evidence and value: impact on DEcisionMaking—the EVIDEM framework and potential applications. *BMC Health Serv Res* 2008;8:270.
- [10] Goetghebeur MM, Wagner M, Khoury H, et al. Combining multicriteria decision analysis, ethics and health technology assessment: applying the EVIDEM decision-making framework to growth hormone for Turner syndrome patients. *Cost Eff Resour Alloc* 2010;8:4.
- [11] Dowie J. The future of HTA is MCDA. 2008. Available from: <http://cafeannalisa.wordpress.com/article/the-future-of-hta-is-mcda-2rxcx200qipsa-14/>. [Accessed April 27, 2012].
- [12] Mitton CR, Donaldson C, Waldner H, Eagle C. The evolution of PBMA: towards a macro-level priority setting framework for health regions. *Health Care Manag Sci* 2003;6:263–9.
- [13] Mitton CR, Donaldson C. Setting priorities and allocating resources in health regions: lessons from a project evaluating program budgeting and marginal analysis (PBMA). *Health Policy* 2003;64:335–48.
- [14] Peacock SJ, Richardson JR, Carter R, Edwards D. Priority setting in health care using multi-attribute utility theory and programme budgeting and marginal analysis (PBMA). *Soc Sci Med* 2007;64:897–910.
- [15] Liberatore MJ, Nydick RL. The analytic hierarchy process in medical and health care decision making: a literature review. *Eur J Oper Res* 2008;189:194–207.
- [16] Vaidya OS, Kumar S. Analytic hierarchy process: an overview of applications. *Eur J Oper Res* 2006;169:1–29.
- [17] Aouni B, Kettani O. Goal programming model: a glorious history and a promising future. *Eur J Oper Res* 2001;133:225–31.
- [18] Claxton K, Briggs A, Buxton MJ, et al. Value based pricing for NHS drugs: an opportunity not to be missed? *BMJ* 2008;336:251–4.
- [19] Towse A. Value based pricing, research and development, and patient access schemes: will the United Kingdom get it right or wrong? *Br J Clin Pharmacol* 2010;70:360–6.
- [20] Eichler HG, Kong SX, Gerth WC, et al. Use of cost-effectiveness analysis in health-care resource allocation decision-making: how are cost-effectiveness thresholds expected to emerge? *Value Health* 2004;7:518–28.
- [21] Dolan P, Tsuchiya A, Wailoo A. NICE's citizen's council: what do we ask them, and how? *Lancet* 2003;362:918–9.
- [22] Dolan P, Shaw R, Tsuchiya A, Williams A. QALY maximisation and people's preferences: a methodological review of the literature. *Health Econ* 2005;14:197–208.
- [23] Sassi F, Archard L, Le GJ. Equity and the economic evaluation of healthcare. *Health Technol Assess* 2001;5:1–138.
- [24] Williams AH, Cookson RA. Equity-efficiency trade-offs in health technology assessment. *Int J Technol Assess Health Care* 2006;22:1–9.
- [25] Wailoo A, Tsuchiya A, McCabe C. Weighting must wait: incorporating equity concerns into cost-effectiveness analysis may take longer than expected. *Pharmacoeconomics* 2009;27:983–9.
- [26] Stolk EA, Poley MJ. Criteria for determining a basic health services package: recent developments in The Netherlands. *Eur J Health Econ* 2005;6:2–7.
- [27] Annemans L, Crott R, De CH, et al. Pricing and reimbursement of pharmaceuticals in Belgium. *Pharmacoeconomics* 1997;11:203–9.
- [28] Sandier S, Paris V, Polton D. *Health Care Systems in Transition: France*. Copenhagen: WHO Regional Office for Europe on behalf of the European Observatory on Health Systems and Policies, 2004.
- [29] Park SE, Lim SH, Choi HW, et al. Evaluation on the first 2 years of the positive list system in South Korea. *Health Policy* 2012;104:32–9.
- [30] Baltussen R, Niessen L. Priority setting of health interventions: the need for multi-criteria decision analysis. *Cost Eff Resour Alloc* 2006;4:14.
- [31] Wilson EC, Peacock SJ, Ruta D. Priority setting in practice: what is the best way to compare costs and benefits? *Health Econ* 2009;18:467–78.
- [32] Baltussen R, Youngkong S, Paolucci F, Niessen L. Multi-criteria decision analysis to prioritize health interventions: capitalizing on first experiences. *Health Policy* 2010;96:262–4.
- [33] Husereau D, Boucher M, Noorani H. Priority setting for health technology assessment at CADTH. *Int J Technol Assess Health Care* 2010;26:341–7.
- [34] Department for Communities and Local Government. *Multicriteria Analysis: A Manual*. London: Department for Communities and Local Government, 2009.
- [35] Nutt DJ, King LA, Phillips LD. Drug harms in the UK: a multicriteria decision analysis. *Lancet* 2010;376:1558–65.
- [36] Mussen F, Salek S, Walker S. A quantitative approach to benefit-risk assessment of medicines, part 1: the development of a new model using multi-criteria decision analysis. *Pharmacoepidemiol Drug Saf* 2007;16(Suppl. 1):S2–15.
- [37] Mussen F, Salek S, Walker S. *Development and Application of a Benefit-Risk Assessment Model Based on Multi-Criteria Decision Analysis*. Chichester, UK: John Wiley & Sons, Ltd., 2009.
- [38] Dolan JG. Shared decision-making—transferring research into practice: the analytic hierarchy process (AHP). *Patient Educ Couns* 2008;73:418–25.
- [39] van Til JA, Renzenbrink GJ, Dolan JG, Ijzerman MJ. The use of the analytic hierarchy process to aid decision making in acquired equinovarus deformity. *Arch Phys Med Rehabil* 2008;89:457–62.
- [40] Dolan JG. Medical decision making using the analytic hierarchy process: choice of initial antimicrobial therapy for acute pyelonephritis. *Med Decis Making* 1989;9:51–6.
- [41] Hummel JM, Snoek GJ, van Til JA, et al. A multicriteria decision analysis of augmentative treatment of upper limbs in persons with tetraplegia. *J Rehabil Res Dev* 2005;42:635–44.
- [42] Ruta D, Mitton C, Bate A, Donaldson C. Programme budgeting and marginal analysis: bridging the divide between doctors and managers. *BMJ* 2005;330:1501–3.
- [43] Mitton C, Donaldson C. Twenty-five years of programme budgeting and marginal analysis in the health sector, 1974–1999. *J Health Serv Res Policy* 2001;6:239–48.
- [44] Posnett J, Street A. Programme budgeting and marginal analysis: an approach to priority setting in need of refinement. *J Health Serv Res Policy* 1996;1:147–53.
- [45] Robinson R. Limits to rationality: economics, economists and priority setting. *Health Policy* 1999;49:13–26.
- [46] Health England. *Prioritising Investments in Preventative Health*. England: Health England, 2009.
- [47] National Institute for Health and Clinical Excellence. *National Institute for Health and Clinical Excellence (NICE) Guide to the Methods of Technology Appraisal*. London: National Institute for Health and Clinical Excellence, 2008. Available from: <http://www.nice.org.uk/>. 2008. [Accessed November 10, 2010].
- [48] Phillips L, Bana e Costa C. Transparent prioritisation, budgeting and resource allocation with multi-criteria decision analysis and decision conferencing. *Ann Oper Res* 2007;154:51–68.
- [49] Schein E. *Process Consultation Revisited: Building the Helping Relationship* (Addison-Wesley Series on Organization Development). Essex, UK: Addison-Wesley Longman, 1999.
- [50] Devlin N, Sussex J. *Incorporating Multiple Criteria in HTA: Methods and Processes*. London: Office of Health Economics, 2011.
- [51] Roy B. The outranking approach and the foundations of ELECTRE methods. *Theory Decis* 1991;31:49–73.

- [52] Figueira J, Greco S, Ehrogott M, et al. *Electre Methods; Multiple Criteria Decision Analysis: State of the Art Surveys*. International Series in Operations Research and Management Science. New York: Springer, 2005.
- [53] Roy B. Ranking and choice in pace of multiple points of view (ELECTRE method). *Revue Francaise D Informatique de Recherche Operationnelle* 1968;8:57–75.
- [54] Roy B, Vanderpooten D. The European school of MCDA: emergence, basic features and current works. *J Multi-Crit Decis Anal* 1996;5:22–38.
- [55] Brans JP, Vincke P. A preference ranking organisation method: (the PROMETHEE method for multiple criteria decision-making). *Manag Sci* 1985;31:647–56.
- [56] Brans JP, Mareschal B. The PROMCALC & GAIA decision support system for multicriteria decision aid. *Decis Support Syst* 1994;12:297–310.
- [57] Ignizio JP. A review of goal programming: a tool for multiobjective analysis. *J Oper Res Soc* 1978;29:1109–19.
- [58] Tamiz M, Jones D, Romero C. Goal programming for decision making: an overview of the current state-of-the-art. *Eur J Oper Res* 1998;111: 569–81.
- [59] Ryan M, Gerard K. Using discrete choice experiments to value health care programmes: current practice and future research reflections. *Appl Health Econ Health Policy* 2003;2:55–64.
- [60] Ryan M. Discrete choice experiments in health care. *BMJ* 2004;328: 360–1.
- [61] Dolan JG. Multi-criteria clinical decision support: a primer on the use of multiple criteria decision making methods to promote evidence-based, patient-centered healthcare. *Patient* 2010;3:229–48.
- [62] Hsia P, Samuel J, Gao J, et al. Formal approach to scenario analysis. *Software, IEEE* 1994;11:33–41.
- [63] Winterfeldt von D, Fischer GW. Multi-attribute utility theory: Models and assessment procedures. In: Wendt D and Viek C, eds. *Utility, Probability and Human Decision Making*. Dordrecht: Reidel, 1975.
- [64] Klir GJ, Yuan B. *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Upper Saddle River, NJ: Prentice-Hall PTR, 1995.
- [65] Tervonen T, Figueira JR. A survey on stochastic multicriteria acceptability analysis methods. *J Multi-Crit Decis Anal* 2008;15: 1–14.
- [66] Hyde K, Maier HR, Colby C. Incorporating uncertainty in the PROMETHEE MCDA method. *J Multi-Crit Decis Anal* 2003;12:245–59.
- [67] Dorini G, Kapelan Z, Azapagic A. Managing uncertainty in multiple-criteria decision making related to sustainability assessment. *Clean Technol Environ Policy* 2011;13:133–9.