

Value-Based Assessment of New Medical Technologies: Towards a Robust Methodological Framework for the Application of Multiple Criteria Decision Analysis in the Context of Health Technology Assessment

Aris Angelis¹ · Panos Kanavos¹

Published online: 6 January 2016

© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract In recent years, multiple criteria decision analysis (MCDA) has emerged as a likely alternative to address shortcomings in health technology assessment (HTA) by offering a more holistic perspective to value assessment and acting as an alternative priority setting tool. In this paper, we argue that MCDA needs to subscribe to robust methodological processes related to the selection of objectives, criteria and attributes in order to be meaningful in the context of healthcare decision making and fulfil its role in value-based assessment (VBA). We propose a methodological process, based on multi-attribute value theory (MAVT) methods comprising five distinct phases, outline the stages involved in each phase and discuss their relevance in the HTA process. Importantly, criteria and attributes need to satisfy a set of desired properties, otherwise the outcome of the analysis can produce spurious results and misleading recommendations. Assuming the methodological process we propose is adhered to, the application of MCDA presents three very distinct advantages to decision makers in the context of HTA and VBA: first, it acts as an instrument for eliciting preferences on the performance of alternative options across a wider set of explicit criteria, leading to a more complete assessment of value; second, it allows the elicitation of preferences across the criteria themselves to reflect differences in their relative importance; and, third, the entire process of preference elicitation can be informed by direct stakeholder engagement, and can

therefore reflect their own preferences. All features are fully transparent and facilitate decision making.

Key Points for Decision Makers

Multiple criteria decision analysis (MCDA) has emerged as a likely alternative approach to economic evaluation in the context of health technology assessment (HTA). However, there is no sufficient methodological guidance on how to design, conduct and implement MCDA as part of HTA, including how to select criteria appropriately.

An MCDA-based methodological framework in the context of HTA could be divided into the phases of problem structuring, model building, model assessment, model appraisal, and action plans. For the analysis to be robust and for decision recommendations to be ultimately meaningful, criteria and attributes should adhere to a number of properties.

The resulting MCDA index score could act as a more encompassing measure of value given that multiple benefit dimensions are incorporated. Consideration of purchasing costs could be used to derive the different options' incremental cost value ratio (ICVR) and contribute to priority setting and resource allocation.

✉ Aris Angelis
a.n.angelis@lse.ac.uk

¹ Department of Social Policy and Medical Technology Research Group, LSE Health, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK

1 Introduction

The use of economic evaluation methods, particularly cost-effectiveness analysis (CEA) and cost-utility analysis (CUA), to assess the incremental benefit of new medical

technologies in relation to the best alternative care has increased considerably over the past 2 decades. In this context, the use of the quality-adjusted life-year (QALY) has been established as the preferred measure of health gain across many settings [1–4]. This is despite its frequent dependence on restrictive assumptions [5], the non-alignment of public versus patients' decision utilities, which would differ from their respective experienced utilities¹ [6], and the reliance on generic tools, such as the EQ-5D, that may not reflect patient experience adequately [7–9].

At the same time, there is increased recognition that economic evaluation has limitations because it does not capture a number of important dimensions of value, and is therefore lacking in comprehensiveness. In partial recognition of that, economic evaluation has recently evolved into a deliberative process across different settings, whereby independent decision-making committees often allow for other dimensions of value to be considered, at least implicitly.

Additionally, there is increasing evidence that decision makers are reluctant to make coverage recommendations on economic evaluation alone [10] and, consequently, 'value' based on economic evaluation results could be informed by additional dimensions of benefit. Recently, decision makers in England and Wales considered additional parameters of benefit on an ad hoc basis [11], highlighting the need to seek a broader and more transparent assessment methodology [12, 13], in the context of value-based pricing [14–17].

Even under such enhanced settings, the decision-making framework often lacks transparency, not least because different stakeholders attach different value judgements to the criteria considered. Consequently, value assessment is not simply a question of what additional benefits to consider and possibly include in the decision-making process, but, importantly, involves how to arrive at a clear process that elicits and accounts for the preferences of different stakeholders in a transparent way. The ongoing debate in the UK on value-based pricing is a testament to these issues [18].

Overall, the lack of comprehensiveness, the emerging ad hoc and non-systematic use of additional dimensions of value and the lack of transparency in making value judgements often lead to inconsistencies in the appraisal process. The consequence is 'unexplained' decision heterogeneity, with important implications for fairness, equity and resource allocation. The development of alternative methodological approaches for value assessment of medical technologies that would overcome the above

limitations could contribute to a more complete framework of value assessment and, in turn, lead to more efficient resource allocation.

In recent years, multiple criteria decision analysis (MCDA) has emerged as a likely alternative to address the current shortcomings of Health Technology Assessment (HTA) based on economic evaluation [19–24]. One of the conclusions of a recent review of MCDA approaches adopted in healthcare was that decision makers display a positive attitude towards its potential to improve decision making [25]. Conceptually, there are three main reasons why MCDA could provide a useful alternative to economic evaluation-based HTA processes. The first relates to the inclusion of a comprehensive list of value dimensions in an explicit manner, beyond what economic evaluation methods currently capture. This enables value assessment to be conducted in an encompassing manner and, in principle, addresses a key limitation of economic evaluation. The second relates to the assignment of quantitative weights across the different evaluation criteria. In doing so, the relative importance of various value dimensions is explicitly incorporated, improving the transparency of the preference-elicitation process. The third is stakeholder participation and the possibility to include all relevant stakeholders in the value-assessment process. This is both insightful—enabling stakeholder views to be heard in a dynamic environment, where all inputs are considered prior to making decisions about coverage—and politically correct, increasing the legitimacy of decision processes, as all stakeholder views are accounted for in an open and transparent way.

Despite the above, the methodological details of MCDA implementation in the context of healthcare decision making have not been sufficiently discussed, and there is no adequate guidance on how MCDA should be conducted in HTA, particularly in relation to which criteria to incorporate and how.

In this paper, we outline a methodological process for the development of a robust MCDA framework and debate its implementation in the context of HTA. In doing so, we provide a broad classification of MCDA methods while also accounting for and building on the classifications proposed in the literature [26–32]. We then focus on value-based methods, specifically MAVT methods, and argue in favour of using these because of their comprehensive nature. Further, we argue that several key principles need to be fulfilled in order for any MCDA framework to be methodologically sound and for the results produced to be robust and policy relevant. These principles apply to the MCDA main phases and stages as well as to the properties that the selected criteria and attributes need to satisfy, while establishing their relevance in the context of HTA and value-based assessment (VBA). We discuss these

¹ "Decision utility" refers to the preference or desire for an outcome that has not occurred in contrast to "experienced utility" which refers to the actual hedonic experience of an outcome [6].

principles and their implications in the context of HTA by drawing on concrete examples. Finally, we discuss a number of practical issues relating to the use of MCDA in HTA and provide a link to policy making.

2 A Methodological Framework Applying MCDA Principles in HTA and Value Assessment of Medical Technologies

2.1 MCDA in the Context of Multi-Attribute Value Theory

MCDA is both an approach and a group of techniques aiming to aid decision making by laying out the problem, objectives and available options in a clear and transparent way. Different MCDA methods exist, with variable degrees of complexity making use of different analytic models. These methods can be broadly categorised by ‘school of thought’, notably (1) value-measurement methods, including (multi-attribute) value theory and utility theory methods, (2) ‘satisficing’ and aspiration level methods, (3) outranking methods, and (4) fuzzy and rough sets methods [26–29]. However, no universal categorisation of MCDA methods exists, and others have proposed groupings that differ from the above [30–32]. Each MCDA method has its own advantages and disadvantages. The choice of method is informed by the type of problem to be addressed, the type of judgements required, the set of axioms employed to support decision making, and the kind of responses needed. Some methods address choice problems, while others address ranking problems or classification and sorting problems.

The methodological process we are proposing in this paper for the context of HTA pertains to the category of value-measurement methods. This is predominantly because of the multiple problems that can be addressed, the simplicity of the judgements required and the relatively limited restrictions imposed by the axioms employed. The value-measurement methods category is widely used in healthcare because of these features. Nevertheless, some aspects (e.g. the MCDA phases and the criteria properties) are applicable across different MCDA methods beyond the value-measurement methods category.

Value-measurement methods usually aim to address ranking or choice problems, ordering a set of alternative options with respect to their performance on a number of objectives or criteria, through the production of overall numerical value scores. A value (or real number) V is associated with the performance of an alternative a , in order to produce an ordering of preferences for all alternatives being considered, while being consistent with the assumptions of complete and transitive preferences. These

methods include linear additive methods, multi-attribute value theory (MAVT) methods for deterministic consequences, and multi-attribute utility theory (MAUT) methods.

We argue in favour of MAVT methods because of their comprehensiveness and methodological robustness [27], as well as their ability to reduce ambiguity and motivational biases. The MAVT methods framework adheres to a number of phases and stages and includes (1) the definition of objectives, (2) the selection of criteria, (3) the scoring of options, and (4) the assignment of weights to the selected criteria.

The choice of technique that will inform parts of the process, including scoring, weighing and aggregation, is an important decision. Under MAVT methods, partial value functions for individual criteria are constructed in the first instance and are subsequently aggregated. Essentially, value functions reflect decision makers’ preferences for different levels of performance on the attribute scale (Fig. 1). Importantly, the assumptions required for the formation of the partial value functions are interlinked with the aggregation type of technique used. In the sections that follow, we present and discuss these fundamental principles in the context of healthcare decision making and use examples to illustrate their application and interpretation.

2.2 MCDA Phases Under MAVT Methods

While the general features of MCDA phases have already been discussed elsewhere [26], the MCDA process could be divided into five distinct phases in the context of HTA; these would be (1) problem structuring, (2) model building, (3) model assessment, (4) model appraisal, and (5) development of action plans (Fig. 2).

Problem structuring involves an understanding of the problem to be addressed. This includes key concerns, envisaged goals, relevant stakeholders that may participate in or contribute to decisions, and identification of uncertainties in terms of a new technology’s clinical evidence and its quality.

The phases of model building, model assessment and model appraisal involve the construction of decision makers’ value judgements within and across the criteria of interest, while being consistent with a set of assumptions, aiming to help decision makers elicit and order their preferences across the alternative options evaluated. For example, if overall survival (OS) is a criterion, then the respective value associated with a range of plausible incremental OS gains (e.g. 3, 6, 9, or more months) is of interest to know and so is the intensity with which stakeholders would prefer certain changes within the attribute range (e.g. an increase in OS from 3 to 6 months could be

Fig. 1 Value function for scoring the performance of alternative options

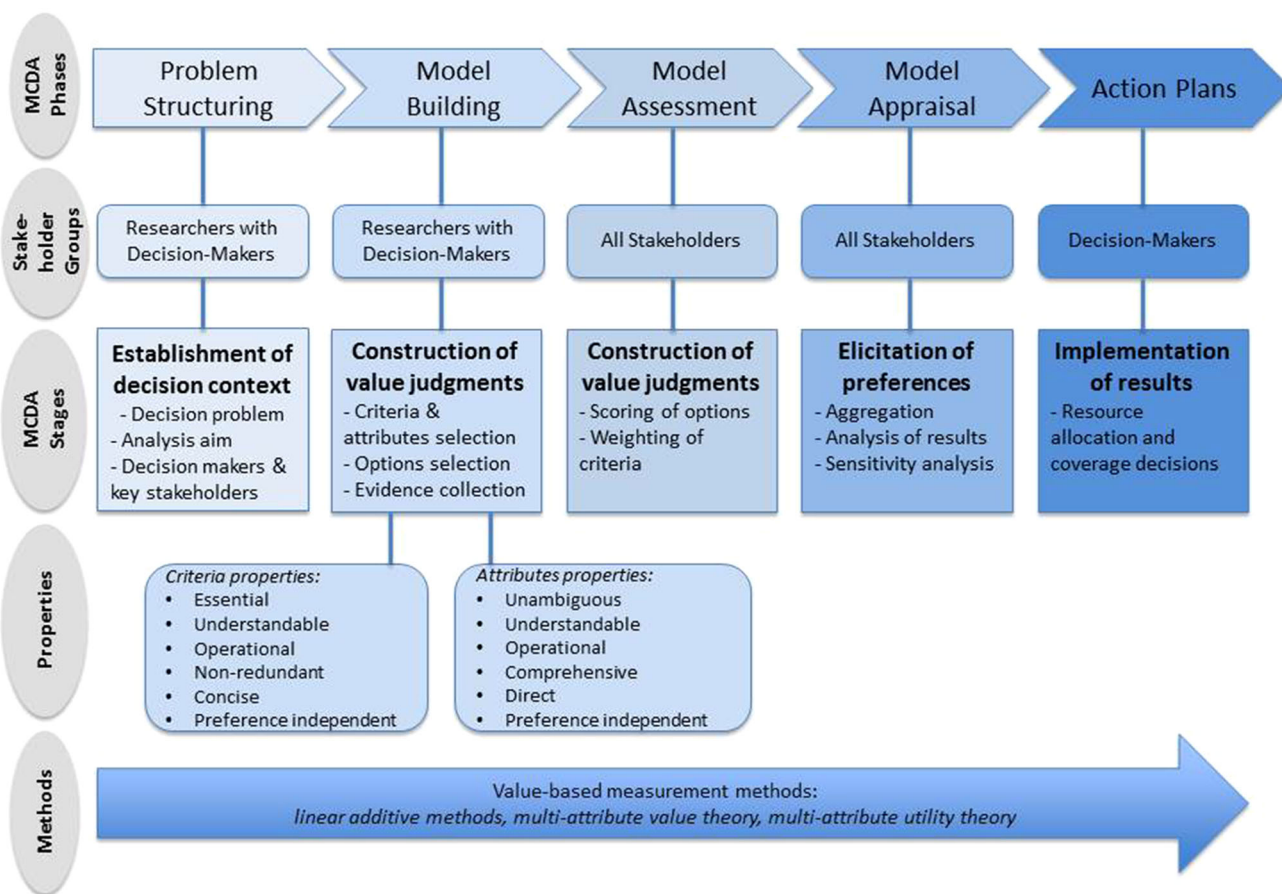
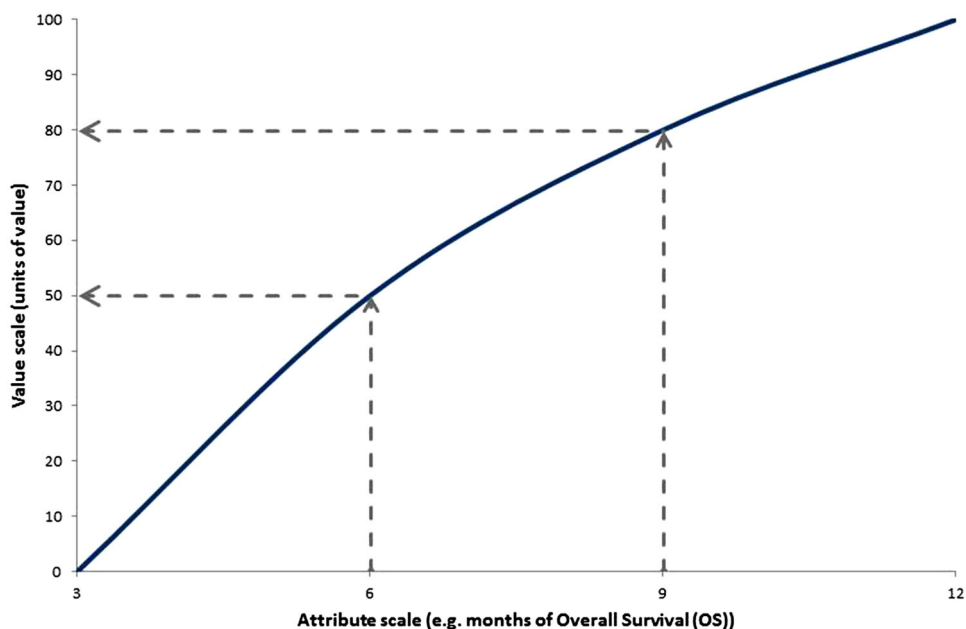


Fig. 2 Multiple criteria decision analysis (MCDA) methodological process in the context of health technology assessment

of greater value to some stakeholders than an increase in OS from 6 to 9 months) (Fig. 1).

Finally, given that the outcome of the analysis needs to inform decision making, action plans need to be shaped involving a clear pathway for result implementation. In the case of HTA, this could involve prioritising resource allocation as part of coverage decisions that take place following the evaluation of new medical technologies.

Although these five phases are presented as part of a linear process, in reality they could be part of an iterative process, moving from a later step back to a previous step before advancing. For example, as part of the model assessment phase, it could become evident that some of the criteria do not possess all the required properties (see Sect. 3), in which case the model should be adapted accordingly as part of the model-building phase.

2.3 MCDA Stages under MAVT Methods

2.3.1 Problem Structuring

Each MCDA phase comprises a number of stages (Fig. 2). Initially, as part of the problem-structuring phase, the decision context needs to be established where the problem under investigation and the aims of the analysis are clearly outlined and defined, and relevant decision makers and other key stakeholders are identified. For example, in the context of VBA of a new technology, the decision problem may be to assess the new technology's benefits and costs from a broader societal perspective relative to other therapeutic alternatives to identify the most valuable treatment for a health system. The decision makers in this context would be payers or insurers (including commissioners of care), whereas healthcare professionals, patients and their carers, technology suppliers and methodology experts, including decision-analysis experts conducting and co-ordinating the MCDA process, would be the relevant stakeholders. The process of identifying the appropriate decision makers and stakeholders would be specific to the country or setting. This particular phase could be conducted by researchers or, alternatively, an HTA agency in settings where such an agency exists.

2.3.2 Model Building

Subsequently, as part of the model-building phase, objectives need to be established and/or relevant criteria identified to reflect decision makers' goals and areas of concern. Additionally, attributes need to be selected to operationalise these criteria and enable their assessment. This involves a deliberative process in order to obtain a good understanding of the decision problem and what decision makers want to achieve (objectives), through

which the values of concern (criteria) will eventually emerge. The assessment takes place based on the selected criteria and attributes. For example, when evaluating a new medical technology relative to an older one, criteria from a number of domains could be selected, such as therapeutic benefit, safety profile, burden of illness, innovation level and socioeconomic impact [22, 33]. In principle, these criteria domains would emerge from decision makers' values of concern; in practice, they could be identified from the literature in combination with semi-structured interviews with decision makers. Quality of evidence, mainly relating to relevance and validity of the available evidence, is another crucial parameter that should be considered. This phase could be carried out by MCDA researchers in collaboration with the decision makers and possibly stakeholders whose value concerns should be considered.

As part of the model-building phase, the alternative options need to be selected, and evidence on their performance across criteria/attributes needs to be identified. For example, the treatment alternatives for a particular disease must be identified and data on expected or observed performance across criteria must be collected, either through secondary research (e.g. from published randomized controlled trial results) or through primary research if data are not available from secondary sources (e.g. clinical or patient opinion). Following the completion of this stage, attribute ranges will be set based on the performance of the alternative treatment options that shall inform the next stages of the process. Depending on the technique used, plausible attribute ranges can be set by taking into account any pre-existing preferences of decision makers in relation to maximum and minimum allowable performance levels on the different criteria. For example, the OS gains of three different treatments could range from 2 to 12 months, and therefore the respective attribute range should be broad enough to include all these gains (i.e. at least from 2 to 12 months). It could also be the case that the decision maker is not willing to consider any treatments offering incremental OS gains of less than 3 months; in this instance, the attribute range could be rescaled and adapted to decision makers' revealed preferences (i.e. to range from a minimum of 3 months upwards), with the treatment option offering 2 months of OS excluded from the analysis.

2.3.3 Model Assessment

In the context of the model-assessment phase, the performance of options against the identified criteria must be assessed (i.e. scoring, which delivers intra-criteria information), and criteria must be weighed according to their relative importance (i.e. weighing, which delivers inter-criteria information), revealing preferences for different

levels of performance within criteria and across different criteria, respectively. In the case of the OS example, a numerical value score would be assigned to the options being evaluated with regards to their performance on OS gains. As part of MAVT methods, the construction of value functions can take place through different techniques (direct rating, indirect, bisection techniques). All require the definition of attribute reference levels that will form the minimum and maximum points of the value scale. Although the two limits of the attribute range are usually assigned a value of 0 and 100, reflecting the minimum and maximum points of the value scale, respectively, other reference points can also be used. Using the OS example, 3 months could be used as the lower reference level and 12 months as the higher reference level, making up the 0 and 100 points of the value scale, respectively. The attribute performance of the options can then be assessed indirectly through the use of the value functions that will convert their performance into value scores (Fig. 1). The process of scoring and weighing completes the construction of value judgements.

A critical aspect in the entire process is the relative importance of the different criteria to decision makers. For this reason, relative weights are assigned to the criteria by directly involving decision makers and stakeholders. For example, in the case of a new drug–indication pair, the importance of the therapeutic benefits vis-à-vis an existing therapeutic alternative (e.g. OS gain and quality-of-life improvement) could be found to be twice as important as its safety impact (e.g. adverse events); therefore, the relative weight of the therapeutic cluster of criteria would be twice as high as the product’s safety profile. These weights should only be viewed as scaling constants or trade-off factors, with no algebraic meaning, assigned to enable comparability across criteria to reflect their relative importance.

Methodologically, and contrary to what has been argued elsewhere [19, 24], we would argue that it is important for the criteria weights to be derived ex-post following the selection of the alternative treatments and therefore the formation of the attribute ranges, rather than ex-ante [34]. Theoretically, this is tantamount to arguing for MAVT models, where the construction of value functions precedes the criteria weights, rather than for direct rating methods, where weights are first attached, based on an ex-ante derivation, and the options are then scored. Conceptually, our preference for the ex-post derivation of weights is justified by the nature of health technologies and the conditions they treat: the relative importance of different criteria and, therefore, their respective weights are context specific and depend on the performance of the alternative options in a given context. By means of an example, let us assume that for two treatments (A and B), weights need to

be established for the same criteria (OS and hepatotoxicity), measured through ‘number of months gained’ and ‘incidence of hepatotoxicity’, respectively. The weight assigned to each criterion is very likely to be different if treatment A and B range between 1 and 10 months (1,10 months) in OS and from 10 to 11 % (10,11 %) in hepatotoxicity, compared to the scenario that they range between 10 and 11 months (10,11) and from 1 to 10 % (1,10 %), respectively.

2.3.4 Model Appraisal

As part of the appraisal phase, scores and weights are combined to create a value index (‘aggregating’). The details of this step may differ according to the type of aggregation model used, to include additive or multiplicative value models depending on the level of preference independence present among criteria. Empirical evidence suggests that errors due to the use of additive value aggregation models are in real settings very small and considerably smaller than the errors associated with the wrong aggregation of partial value functions that can possibly result from the incorrect application of more advanced models [26].

Overall, the individual criteria scores and their respective weights are combined to produce weighted scores and are summed to arrive at an overall value score for each treatment option. In combination with sensitivity analysis, results are examined to determine the robustness of the results obtained. The outcome of this process is a ranking of all treatment options based on their respective value scores. Decision makers can use this evidence to make resource-allocation decisions. Throughout the MCDA stages, including scoring and weighing, the participating stakeholders are able to interact to exchange views, reach consensus or simply provide their individual preferences [26]. To that end, they can compare their individual views and preferences, they can aggregate such preferences by voting to reach consensus, or they can share commonly defined modelling and judgement elements after joint discussion.

2.4 MCDA Techniques Using MAVT Methods

Several MCDA techniques are available with regards to scoring, weighing and aggregating. These techniques mainly relate to the value judgement and preference-elicitation processes, and the choice of technique depends on the particular type of method adopted [29–32].

As part of MAVT methods, the value functions based on which options are scored can be constructed using different options: (1) direct rating techniques, (2) indirect techniques, and (3) indifference or bisection techniques [26,

27]. Direct rating techniques involve decisions around the form of the value function and whether they increase monotonically (highest attribute level is the most preferred), decrease monotonically (lowest attribute level is the most preferred), or range non-monotonically (an intermediate attribute level is the most preferred).

Indirect techniques generally assume a monotonic function and involve a series of questions aiming to uncover decision makers' preferences by considering differences in the attribute scale and their relation to the value scale. Indifference techniques explore the magnitude of increments in the attribute scale that correspond to equal units in the value (preference) scale. Finally, bisection techniques explore the estimation of points on the attribute scale that serve as midpoints on the value (preference) scale.

We would argue in favour of indirect elicitation techniques because of their comprehensiveness and unbiased nature. This is mainly because decision-makers' preferences are first elicited for the complete attribute range, and options are then scored indirectly using the attributes' emerging value functions to convert the performance of the options into value scores, essentially by not revealing any information about the identity of the respective options at any point during the process.

An example of such an indirect technique is MACBETH (Measuring Attractiveness by a Categorical Based Evaluation Technique), a convenient indirect approach to elicit value functions by only requiring qualitative judgements about the difference of value between different pairs of attribute levels [35]. It uses seven semantic categories to distinguish between the value or attractiveness of different attribute levels, ranging between "no difference in value" and "extreme difference in value". Overall, it builds a quantitative model of values based on qualitative (verbal) difference judgements, and by analyzing judgmental inconsistencies, it facilitates the move from ordinal preference modelling to cardinal preference modelling.

Once criteria have been scored and value functions have been derived, criteria weights can be elicited, usually through a swing weighting technique. Finally, criteria scores and weights are combined, usually through an additive aggregation approach.

3 Model-Building and the Construction of a Value Tree: Properties to Ensure a Robust MCDA Model

Model building is one of the most important MCDA phases. Establishing objectives and defining the actual criteria and attributes are critical stages in this context because

they form the foundation of MCDA. For the analysis to be robust and, ultimately, meaningful, we outline a number of properties to which criteria and attributes should adhere.

3.1 Objectives, Criteria and Attributes in the Context of Model Building

Depending on the decision problem under consideration, the term 'objective' or 'criterion' may be preferred over the other, both representing key factors that form the basis of the analysis. The main difference between the two is that 'objectives' usually reflect a direction of preference, whereas 'criteria' do not. Objectives and criteria may be further decomposed into sub-objectives and sub-criteria; structuring all objectives and/or criteria in the form of a tree offers an organised overview of the values under consideration. This is known as a value tree. The quantitative or qualitative performance measures associated with criteria or objectives are known as 'attributes'. Attributes operationalize the use of criteria and objectives by measuring the extent to which criteria or objectives are achieved. For example, in the context of a new cancer treatment, an objective for decision makers could be to 'maximise life expectancy'; 'overall survival' could act as a criterion, while 'median number of months from randomisation to death' could be the relevant attribute (Fig. 3).

It is not uncommon for a criterion to require more than a single attribute to be measured adequately. For example, in the case of 'tolerability' as part of a new drug's safety profile, decision makers could benchmark against the 'proportion of patients discontinuing the treatment' as well as the 'proportion of patients interrupting treatment or reducing the dose due to adverse events'. Other examples of value tree hierarchies—made up of criteria and attributes—together with their respective data sources are shown in Fig. 3.

Depending on the type of decision problem, the selection of objectives, criteria and attributes can either precede or follow the identification of the alternative options (Table 1) [26, 36, 37]. In the context of 'value-focused thinking', objectives and criteria are selected prior to specifying or assessing the alternative options, thus being part of a top-down approach for structuring a value tree according to which overall objectives or criteria are decomposed into sub-objectives or sub-criteria [36]. Alternatively, in accordance with the more traditional 'alternative focused thinking', a bottom-up approach can be implemented whereby objectives and criteria emerge following the comparison of the options, based on distinguishable attributes that differentiate them [37].

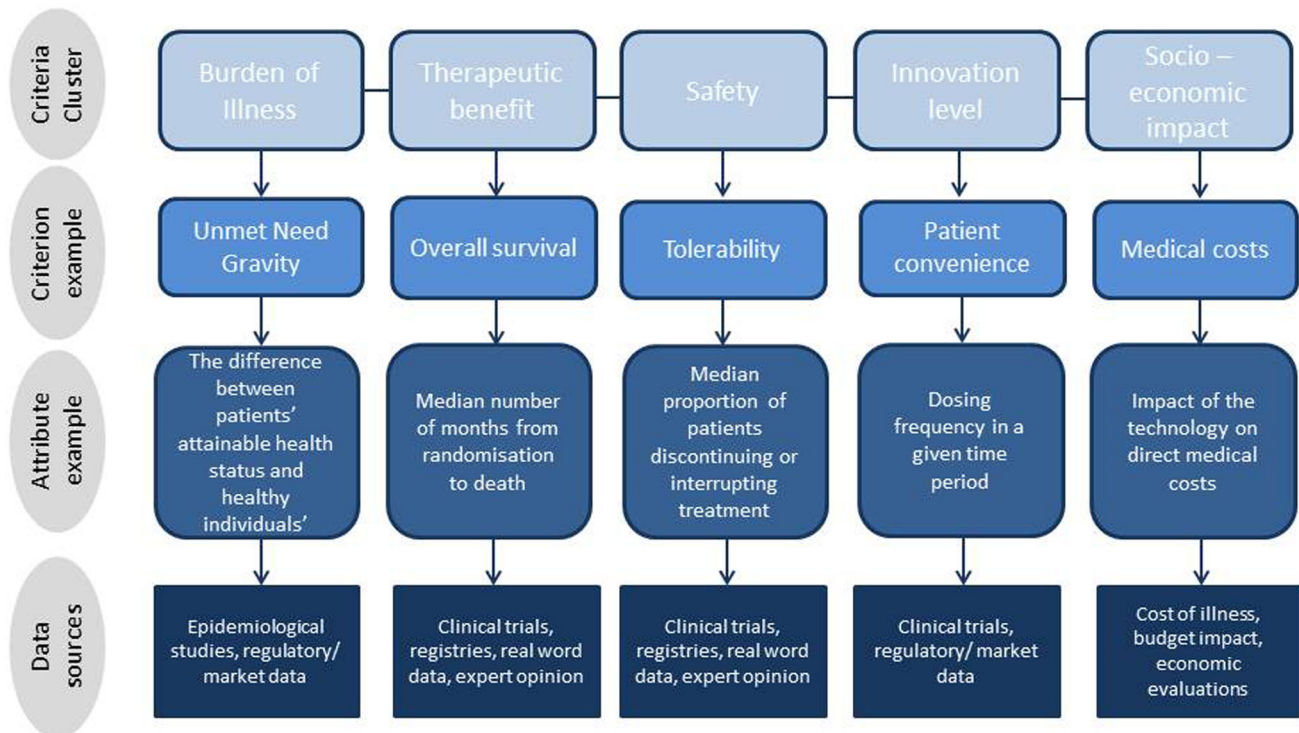


Fig. 3 Value tree hierarchies and data sources using a multiple criteria decision analysis (MCDA) framework for value assessment

Table 1 Different approaches for selecting objectives and criteria

Approach	'Value-focused thinking' [36]	'Alternative-focused thinking' [37]	'Value-alternative hybrid thinking'
Description	Objectives and criteria selected first prior to the identification or assessment of the alternative options	Options first compared so that objectives and criteria can emerge based on their attributes	Generic set of objectives and criteria created first, which then become adapted for the particular decision problem
Value tree formation	Top-down approach	Bottom-up approach	Top-down followed by bottom-up

In the context of HTA, a 'value-alternative hybrid thinking' logic that contains elements from both approaches could be adopted. Decision makers could have a generic set of predetermined objectives and criteria reflecting their values of concern in a top-down approach. These could then be adapted for the purposes of the decision-making problem in a bottom-up approach. Thus, the general values of concern would be tailor made in a dynamic manner to better assess the differences of alternative treatments being compared. For example, decision makers' concerns could normally include the existence of any contraindications or warnings and precautions associated with a drug for the indicated patient population of interest. However, it is possible for all alternative treatments evaluated for a particular disease to have no contraindications and to have identical minor warnings and precautions for use. These criteria could therefore be excluded from that particular assessment in order to be concise.

3.2 Key Criteria Properties

In order for the analysis to provide the highest possible insight and to enhance its actual value to decision makers, both criteria and attributes need to adhere to a number of key properties [36–39]. If they do not, the results obtained through scoring and weighing could be spurious and, therefore, meaningless for decision making. First, objectives or criteria need to be essential, in that all necessary objectives of the decision problem should be considered, and all the critical values under consideration should be included through the incorporation of the respective criteria. In the context of a value tree, all therapeutic, safety, burden of illness, innovation and socioeconomic criteria should be included in the model. Second, criteria need to be understandable, so that all participants in the decision-making process have a clear understanding of them and their implications. Third, criteria need to be operational; namely, the performance of the options against the criteria

should be measurable. Fourth, it is crucial that criteria are non-redundant, i.e. there should be no overlap or double counting between the different criteria, otherwise the elicited criteria weights would not be accurate and, consequently the overall results would be misleading. Finally, criteria need to be concise and only the smallest set that can adequately capture the decision problem should be used, striving for simplicity and parsimony, rather than complexity.

The aggregation stage is very important because it produces the overall value scores of the alternative options. In order to enable the use of simple aggregation rules (e.g. additive value models, where scores and weights of the different individual criteria are multiplied and then added altogether in a weighted average manner), preference independence between the different criteria needs to be upheld [26]. Preference independence is a key property; it implies that an option's value score on a criterion can be elicited independently of the knowledge of the option's performance on the remaining criteria. It should be noted that preference independence is not the same as statistical independence; two different criteria attributes could be statistically dependent but at the same time preference independent and vice versa. If this requirement is not observed, an additive aggregation function should not be employed unless the criteria are restructured to combine non-preference-independent criteria into a single criterion.

If preference independence cannot be satisfied, then more complex aggregation rules (e.g. multiplicative preference functions) would have to be applied to combine scores with weights. Simpler aggregation rules are preferred over more complicated rules in most cases, mainly because they are simple and easy for decision makers to comprehend.

3.3 Key Attribute Properties

For selected attributes to be adequate or meaningful, sufficient properties require them to be unambiguous (in that a clear relationship should exist between the consequences of an option and the levels of attribute used to describe these consequences), comprehensive (the attribute levels should cover the full range of consequences), direct (the attribute levels should describe the consequences of alternative options as directly as possible), operational (information required for the attributes should be collectible in practice and value trade-offs—between the objectives or criteria—can be made), and understandable (the consequences and value trade-offs can be readily understood and communicated across the decision makers and other stakeholders by using the attribute) [39]. For an additive value model to be used, attributes should be preference independent.

A suggested systematic methodology to maximise the probability of selecting the best possible attributes initially involves an aim for a single natural attribute, namely one that is in general use and has a common interpretation measuring directly the degree to which an objective or a criterion is met. If no such single attribute is appropriate then a set (i.e. more than one) of natural attributes should be considered that adequately describe objective/criteria consequences. If this is not possible, exploration of 'constructing' attributes that directly measure consequences should be attempted. Such attributes are explicitly developed to measure directly the achievement of an objective. A proxy attribute, i.e. a less informative attribute that indirectly measures a criterion of concern, should be selected only after careful consideration and following the elimination of constructed attributes [39].

4 From Methodological Robustness to Practical Relevance in the HTA Context

Very often, decision makers and even decision-analysis researchers applying MCDA do not pay sufficient attention to the theoretical foundations of MCDA and the different set of properties that the MCDA models need to possess. Recent evidence has shown that only one healthcare MCDA study explained that criteria were defined to meet MCDA requirements such as avoiding double counting [40], with others acknowledging as a concern the fact that MCDA responders might not understand some of the attributes being used [41], possibly because of difficulties in interpreting the meaning of the respective attribute performance [42].

Taking into consideration that MCDA in itself posits a departure from currently used HTA techniques, the application of an MCDA approach and its principles in the context of HTA requires careful reflection on a number of fronts. First, it is important to clarify whose preferences to consider. Assuming that an HTA agency acts as a proxy decision maker, then it would be appropriate to adopt the perspective of the respective HTA agency. For example, if the decision context is England, France, or Sweden, it would be reasonable to adopt the perspective of the National Institute for Health and Care Excellence (NICE), the Haute Autorite de Sante (HAS) and the Dental and Pharmaceutical Benefits Board (TLV), respectively. Consequently, any social judgements individual HTA agencies adopt, including the participants and their preferences, will need to reflect the particularities of each setting. As different countries or settings are likely to have different priorities and objectives, the analysis should be tailor-made to their needs. Alternatively, if the adoption of such an existing perspective is not possible, some formal

stakeholder analysis could be used to identify the key players that should be involved [43].

A second but related issue is how to combine the preferences of individual stakeholders. Ideally, a consensus approach should be aimed for, through which a single agreed value judgement (i.e. score, weight) would be derived. Alternatively, if mathematical aggregation is used, the median of responders' preferences could be used, especially in settings where motivational biases from some stakeholders exist (i.e. strongly against vs. strongly in favour). In any case, the complete range of value judgements should be recorded and used for sensitivity analysis, where the impact of different scores and weights on the options' total value scores would be tested.

A third practical issue relates to the evidence requirements and their availability. While a criticism of MCDA has been that it requires more evidence than standard HTA approaches to populate the criteria, in practice, similar levels of evidence required for standard HTA approaches could be used in the context of an MCDA model. Even if certain items of information may not be available, they could be readily collectable, or at least able to be proxied through expert opinion.

The consistency of results would be a fourth important issue subject to criticism. Would the results be consistent within the same setting or could inconsistencies act as an obstacle to homogeneous decisions? If the value analysis is evidence based, the results are likely to be different if the participants whose preferences are considered have different value judgements (i.e. different value functions and weights). This highlights the importance of the context wherein MCDA methods should be applied. Indeed, they should act mainly as decision-aiding tools to support priority setting and resource allocation conducted by the decision maker.

5 'Incremental' Versus 'Clean Slate' MCDA Approaches and Link to Policy Making

The application of MCDA in current HTA practices has been criticised partly because criteria should be perceived as attributes of benefit and the fact that cost and uncertainty or quality of evidence cannot be accounted for as benefits [44]. In turn, costs can be considered by incorporating the 'impact on costs' as a criterion, other than the purchasing cost of the treatment itself, which is essentially looking at savings or increased outlays. Poor quality of evidence could be addressed through the incorporation of penalty functions that may be added when significant uncertainty exists, reducing the performance scores of relevant options. For example, if the clinical data relating to an OS gain are regarded as

highly uncertain for any reasons relating to the external and/or internal validity of the clinical trial/data, then the performance score of the observed OS gain for the particular treatment could be reduced by a significant factor, e.g. 25–50 %, based on expert opinion. As identified through a recent review, a number of other formal approaches also exist to quantify and incorporate uncertainty when conducting MCDA for healthcare decisions, the most commonly used being fuzzy set theory, probabilistic sensitivity analysis, deterministic sensitivity analysis, Bayesian framework and grey theory [45].

Another consideration relates to the appropriate process of eliciting weights, and the argument that if they are to emerge during the decision-making process it will prove difficult to achieve predictability, consistency and accountability, but also that scientific and social value judgements might become mixed and prone to strategic behaviour, with pre-specified weights being the way forward [44]. Indeed, producing global weights that are applicable across all decision contexts would seem a very challenging task; however, weights elicited ex-ante would be hardly accurate in capturing the precise trade-offs under consideration for the reasons discussed in Sect. 2.3. By contrast, eliciting weights ex-post would be more reflective of decision-maker preferences and less susceptible to strategic manipulation; however, their application would be mostly restricted at a local decision context. A further source of criticism stems from the question of what attributes of benefit are lost due to additional cost and whether these attributes of benefit foregone can be accounted for by including the incremental cost-effectiveness ratios (ICERs) as criteria [44]. Given that cost effectiveness and cost are not attributes of benefit, one would need to know what additional costs are required to improve the composite measure of benefit and what attributes of benefit will be given up as a consequence of costs.

Some of the above criticisms arise chiefly when MCDA is applied as a 'supplementary' approach to CEA to adjust the ICER by incorporating additional parameters of value. Instead, some of the above criticisms may be overcome by using a 'pure' MCDA approach to derive value without the use of CEA (also described as 'incremental' and 'clean slate' approaches, respectively [20]). In any case, further research would be needed to fully address some of the remaining practical limitations.

The aggregate metric of value emerging from the MCDA process (the value index) is more encompassing in nature, as multiple evaluation criteria are incorporated in the analysis. By adopting this value index metric as the benefit component and incorporating the purchasing cost of the different options, the cost per incremental MCDA value unit gained, i.e. the incremental cost value ratio(s) (ICVR),

could act as the basis of allocating resources in a way comparable to that of an ICER; for instance, options with lower ICVRs would be interpreted as more valuable and could be prioritised versus options with higher ICVRs. Based on this approach, issues relating to the definition of efficiency through the establishment of thresholds that reflect opportunity costs would still need to be addressed; however, they lie outside the scope of this article.

The resulting value index scores would be context specific, reflecting stakeholder preferences: the value index incorporates value judgements and preferences for a set of options based on a group of criteria, all of which can be informed through stakeholder input. Unless identical value judgements and preferences are assumed for the same group of criteria, a value score for an option in one setting could be different from that in another setting. The MCDA process, as proposed in this paper, respects stakeholder preferences in individual settings, whilst reducing decision heterogeneity by introducing clarity, objectivity and greater transparency about the criteria based on which decisions can be shaped.

6 Conclusion

We have proposed a methodological framework outlining the use of MCDA in the context of HTA based on MAVT methods and the respective phases and stages of such a process. Although a variety of MCDA techniques exist, it is likely that the most important stages that act as the foundations to the analysis are the establishment of objectives and the definition of criteria and attributes. We have focused on best practice requirements, as reflected through the appropriate properties needed for criteria and attribute selection, all of which feed into the model-building phase.

Compared with economic evaluation techniques, such as CEA, HTA through MCDA is found to have a number of important advantages. These include the multiplicity of criteria that can be used to assess value, the explicit weights that are assigned to reflect differences in the relative importance of the criteria, the extensive stakeholder engagement across all stages and the transparent nature of the MCDA process, leading to a rounded, flexible, encompassing and transparent approach to value assessment and appraisal. Finally, because of the way MCDA is structured, it facilitates a decision-support system, incorporating important trade-offs as part of the assessment process, whereas in traditional CEA such trade-offs may be considered on an ad hoc basis as part of the appraisal process to inform decision making. Because of its characteristics, HTA through MCDA could be a reasonable

resource-allocation tool that, among others, incorporates a more holistic approach to value.

Acknowledgments The authors are grateful to Gilberto Montibeller and two anonymous referees for providing valuable feedback. Thanks are also due to Lise Rochaix, Karl Arnberg, Thomas Wilkinson, the participants at the HTAi conference in Washington DC (June 2014), and the Advance-HTA dissemination workshops in Warsaw (September 2014) and Mexico City (November 2014) for helpful comments and suggestions on earlier versions of the paper. All outstanding errors are our own. Both authors contributed to all aspects of the paper.

Compliance with Ethical Standards

Funding This study has been prepared under the auspices of Advance-HTA, a research project that has received funding from the European Commission, DG Research (Grant agreement no: 305983). The views represented in the paper do not necessarily reflect the views of the European Commission.

Conflict of interest Aris Angelis has no conflicts of interest to declare that are directly relevant to the content of this article. Panos Kanavos has no conflicts of interest to declare that are directly relevant to the content of this article.

Open Access This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Dolan P. The measurement of health-related quality of life for use in resource allocation decisions in health care. In: Culyer AJ, Newhouse JP, editors. *Handbook of health economics*. Amsterdam: North Holland; 2000.
2. Drummond M. Basic types of economic evaluation. In: Drummond MF, Sculpher MJ, Torrance GW, et al, editors. *Methods for the economic evaluation of health care programmes*. Oxford University Press; 2005.
3. Drummond M, Brixner D, Gold M, et al. Toward a consensus on the QALY. *Value Health*. 2009;12(Suppl 1):S31–5.
4. Weinstein MC, Stason WB. Foundations of cost-effectiveness analysis for health and medical practices. *N Engl J Med*. 1977;296:716–21.
5. Kahneman D. New challenges to the rationality assumption. *Leg Theory*. 1997;3:105–24.
6. Dolan P, Kahneman D. Interpretations of utility and their implications of the valuation of health. *Econ J*. 2008;118:215–34.
7. Normand C. Measuring outcomes in palliative care: limitations of QALYs and the road to PaLYs. *J Pain Symptom Manage*. 2009;38:27–31.
8. Garau M, Shah KK, Mason AR, et al. Using QALYs in cancer: a review of the methodological limitations. *Pharmacoeconomics*. 2011;29:673–85.
9. Tordrup D, Mossman J, Kanavos P. Responsiveness of the EQ-5D to clinical change: in the patient experience adequately represented? *Int J Technol Assess Health Care*. 2014;30:10–9.

10. Zwart-van Rijkom J, Leufkens H, Busschbach J, Broekmans A, Rutten F. Differences in attitudes, knowledge and use of economic evaluations in decision-making in The Netherlands. *Pharmacoeconomics*. 2000;18(2):149–160.
11. National Institute for Health and Clinical Excellence. Appraising life-extending, end of life treatments. London: NICE; 2009.
12. Department of Health. A new value-based approach to the pricing of branded medicines: a consultation. London: Department of Health; 2010.
13. Department of Health. A new value-based approach to the pricing of branded medicines: government response to consultation. London: Department of Health; 2011.
14. Claxton K. OFT, VBP: QED? *Health Econ*. 2007;16:545–58.
15. Office of Fair Trading. The pharmaceutical price regulation scheme: an OFT market study. London: OFT; 2007.
16. Towse A. If it ain't broke, don't price fix it: the OFT and the PPRS. *Health Econ*. 2007;16:653–65.
17. Kanavos P, Manning J, Taylor D, et al. Implementing value-based pricing for pharmaceuticals in the UK. London: 2020health; 2010.
18. National Institute for Health and Care Excellence. Consultation paper: value based assessment of health technologies. London: NICE; 2014.
19. Goetghebeur M, 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.
20. Devlin NJ, Sussex J. Incorporating multiple criteria in HTA: methods and processes. Lond: Office of Health Economics; 2011.
21. Thokala P. Multiple criteria decision analysis for health technology assessment: report by the decision support unit, School of Health and Related Research. Sheffield: University of Sheffield; 2011.
22. Kanavos P, Angelis A. Multiple criteria decision analysis for value based assessment of new medical technologies: a conceptual framework. The LSE Health Working Paper Series in Health Policy and Economics. London: London School of Economics and Political Science; 2013.
23. Sussex J, Rollet P, Garau M, et al. Multi-criteria decision analysis to value orphan medicines. Office of Health Economics Research Papers. London: OHE; 2013.
24. Radaelli G, Lettieri E, Masella C, Merlino L, Strada A, Tringali M. Implementation of EunetHTA core model[®] in Lombardia: the VTS framework. *Int J Technol Assess Health Care*. 2014;30(1):105–12.
25. Marsh K, Lanitis T, Neasham D, Orfanos P, Caro J. Assessing the value of healthcare interventions using multi-criteria decision analysis: a review of the literature. *Pharmacoeconomics*. 2014;32:345–65.
26. Belton V, Stewart TJ. Multiple criteria decision analysis: an integrated approach. Dordrecht: Kluwer Academic Publishers; 2002.
27. Von Winterfeldt D, Edwards W. Decision analysis and behavioral research. Cambridge: Cambridge University Press; 1986.
28. Hammond JS, Keeney RL, Raiffa H. smart choices: a practical guide to making better decisions. Cambridge: Harvard University Press; 1999.
29. Department for Communities and Local Government. Multi-criteria analysis: a manual. London: Communities and Local Government Publications; 2009.
30. 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(4):229–48.
31. Thokala P, Duenas A. Multiple criteria decision analysis for health technology assessment. *Value Health*. 2012;1171–1181.
32. Diaby V, Goeree R. How to use multi-criteria decision analysis methods for reimbursement decision-making in health care: a step-by-step guide. *Expert Rev Pharmacoecon Outcomes Res*. 2014;14(1):81–99.
33. Angelis A, Kanavos P. Using MCDA to assess value of new medical technologies. LSE Health Working Paper Series in Health Policy and Economics. London: London School of Economics and Political Science; 2013.
34. Keeney RL. Common mistakes in making value trade-offs. *Oper Res*. 2002;50(6):935–45.
35. Bana e Costa CA, De Corte JM, Vansnick JC. MACBETH. *Int J Inform Technol Decis Mak*. 2012;11(02):359–87. doi:10.1142/S0219622012400068.
36. Keeney RL. Value focused thinking: a path to creative decision making. Cambridge: Harvard University Press; 1992.
37. Franco LA, Montibeller G. Problem structuring for multicriteria decision analysis interventions. Wiley Encyclopedia of Operations Research and Management Science. Hoboken: Wiley; 2010.
38. Keeney RL, Raiffa H. Decisions with multiple objectives: preferences and value trade-offs. Cambridge: Cambridge University Press; 1993.
39. Keeney RL, Gregory RS. Selecting attributes to measure the achievement of objectives. *Oper Res*. 2005;53:1–11.
40. Hummel MJM, Volz F, Van Manen JG, Danner M, Dintsios CM, Ijzerman MJ, et al. Using the analytic hierarchy process to elicit patient preferences: prioritizing multiple outcome measures of antidepressant drug treatment. *Patient*. 2012;5(4):225–37.
41. Youngkong S, Baltussen R, Tantivess S, Mohara A, Teerawattananon Y. Multicriteria decision analysis for including health interventions in the universal health coverage benefit package in Thailand. *Value Health*. 2012;15(6):961–70.
42. Wilson E, Sussex J, Macleod C, Fordham R. Prioritizing health technologies in a Primary Care Trust. *J Health Serv Res Policy*. 2007;12(2):80–5.
43. Freeman RE. Strategic management: a stakeholder approach. Boston: Pitman; 1984.
44. Claxton K. Should multi-criteria decision analysis (MCDA) replace cost effectiveness analysis (CEA) for evaluation of health care coverage decisions? In: ISPOR 16th Annual European Congress. Dublin; 2013.
45. Broekhuizen H, Groothuis-Oudshoorn CGM, van Til JA, Hummel JM, Ijzerman MJ. A review and classification of approaches for dealing with uncertainty in multi-criteria decision analysis for healthcare decisions. *Pharmacoeconomics*. 2015;33:445–55.