

Evaluating the Benefits of New Drugs in Health Technology Assessment Using Multiple Criteria Decision Analysis: A Case Study on Metastatic Prostate Cancer With the Dental and Pharmaceuticals Benefits Agency (TLV) in Sweden

MDM Policy & Practice

1–17

© The Author(s) 2018

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/2381468318796218

journals.sagepub.com/home/mdm

Aris Angelis 

Abstract

Background. Multiple criteria decision analysis (MCDA) has been identified as a prospective methodology for assisting decision makers in evaluating the benefits of new medicines in health technology assessment (HTA); however, limited empirical evidence exists from real-world applications. **Objective.** To test in practice a recently developed MCDA methodological framework for HTA, the Advance Value Framework, in a proof-of-concept case study with decision makers. **Methods.** A multi-attribute value theory methodology was adopted applying the MACBETH questioning protocol through a facilitated decision-analysis modelling approach as part of a decision conference with four experts. **Settings.** The remit of the Swedish Dental and Pharmaceutical Benefits Agency (Tandvårds- och läkemedelsförmånsverket [TLV]) was adopted but in addition supplementary value dimensions were considered. **Patients.** Metastatic castrate-resistant prostate cancer patients were considered having received prior chemotherapy. **Interventions.** Abiraterone, cabazitaxel, and enzalutamide were evaluated as third-line treatments. **Measurements.** Participants' value preferences were elicited involving criteria selection, options scoring, criteria weighting, and their aggregation. **Results.** Eight criteria attributes were finally included in the model relating to therapeutic impact, safety profile, socioeconomic impact, and innovation level with relative importance weights 44.5%, 33.3%, 14.8%, and 7.4% per cluster, respectively. Enzalutamide scored the highest overall weighted preference value score, followed by abiraterone and cabazitaxel. Dividing treatments' overall weighted preference value scores by their costs derived "costs per unit of value" for ranking the treatments based on value-for-money grounds. **Limitations.** Study limitations included lack of comparative clinical effects across treatments and the small sample of participants. **Conclusion.** The Advance Value Framework has the prospects of facilitating the evaluation process in HTA and health care decision making; additional research is recommended to address technical challenges and optimize the use of MCDA for policy making.

Keywords

Advance Value Framework, case study, Dental and Pharmaceuticals Benefits Agency (TLV), health technology assessment (HTA), medicines, multiple criteria decision analysis (MCDA), prostate cancer, Sweden, value assessment

Date received: November 17, 2017; accepted: July 17, 2018

In recent years, the assessment and appraisal of new medicines by health technology assessment (HTA) agencies and payer bodies for the purpose of pricing and reimbursement

Corresponding Author

Aris Angelis, Medical Technology Research Group, LSE Health and Department of Health Policy, London School of Economics and Political Science, Portugal St, London WC2A 2AE, UK; telephone +44 207 955 6842 (a.n.angelis@lse.ac.uk).



This Creative Commons Non Commercial CC-BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<http://www.creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

has received considerable attention. At times, this has emerged as a consequence of negative coverage decisions for some medicines, often related to treatments for high-burden diseases associated with significant mortality and morbidity. As a consequence, intense disputes for the allocation of resources often arise usually involving patients, clinicians, and pharmaceutical manufacturers on one hand and HTA agencies, commissioners of care, and public decision makers on the other.

To a large extent this conflict can be attributed to current methodologies used for the evaluation of new medicines as they cannot adequately capture the different notions of value evident given the existence of diverse perspectives.¹ For example, despite the fact that the QALY (quality-adjusted life-year) metric has managed to successfully combine the two vital dimensions of health benefit relating to life expectancy and health-related quality of life (HRQoL), its sole and inflexible adoption as part of economic evaluations in HTA can at times be regarded as blunt and insufficient, not least because it does not reflect other important value dimensions of treatments.^{2,3} This lack of comprehensiveness has led decision makers to use additional parameters of value in an informal and nonsystematic manner, often diminishing transparency in the evaluation procedures,⁴ with emerging inconsistencies in coverage threatening the credibility of the decision outcomes and their acceptance or perceived fairness from the public.

More recently, a growing number of resource-conscious health care professionals have started to publicly oppose the use of highly expensive new drugs in clinical practice that are perceived to bring marginal added clinical benefit on the grounds of poor value for money. Escalating drug prices have catalyzed the generation of numerous “value frameworks” with the aim to inform payers, clinicians, and patients around the assessment process of new medicines for the purpose of coverage and treatment selection decisions.^{5–8} Although this is an important step toward a more inclusive value-based assessment approach and should be supported,⁹ aspects of these frameworks might be based on weak or

atheoretical methodologies and could potentially result in misleading recommendations or decisions.¹⁰

Research and development of alternative methodologies for assessing and appraising the value of new medicines could overcome such limitations, contributing toward a more complete framework for measuring value and making resource allocation and treatment decisions. Multiple criteria decision analysis (MCDA) has recently surfaced as an alternative method not only to traditional economic evaluation techniques with the prospects of addressing some of their limitations in HTA^{11–18} but also for eliciting patient preferences and facilitating treatment selection.^{19–21} In any case, empirical evidence from real-world applications with HTA decision makers are very limited.

This is a proof-of-concept case study in an HTA context applying a recently developed value framework, the Advance Value Framework (AVF),^{22,23} for assessing the value of metastatic castrate-resistant prostate cancer (mCRPC) treatment options following second-line chemotherapy. This corresponds to the first of a series of case studies that test in practice the AVF by engaging with and eliciting the preferences of decision makers from national HTA agencies and payer bodies in Europe. It forms continuation of pilot work from an earlier case study on the assessment of drugs for metastatic colorectal cancer using value judgements from a range of stakeholders in the English setting.²⁴ A facilitated workshop was organized taking the form of a decision conference,²⁵ adopting a facilitated decision analysis modelling approach,^{26,27} and involving an experts panel of assessors from the Swedish Dental and Pharmaceutical Benefits Agency (Tandvårds- och läkemedelsförmånsverket [TLV]). Metastatic prostate cancer was chosen as a case study topic mainly because of its high disease severity and the availability of several expensive treatment options that highlight a clear decision problem. This indication has been a popular appraisal topic by numerous HTA agencies, including TLV in Sweden and the National Institute for Health and Care Excellence (NICE) in England.

The adaptation of the AVF for the metastatic prostate cancer setting is described in the Methods section, the rankings of the treatment options based on their value are shown in the Results section, with limitations of the methodology and challenges of MCDA applications in HTA presented in the Discussion section.

Methods

Methodological Framework

The AVF was applied for assessing the value of a set of drug treatments, similarly to a previous application for

Medical Technology Research Group, LSE Health and Department of Health Policy, London School of Economics and Political Science, London, UK. The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. Financial support for this study was provided entirely by a grant from the European Commission, DG Research (Advance HTA project, Grant Number 305983). The funding agreement ensured the author's independence in designing the study, interpreting the data, writing, and publishing the report. The views represented in the article do not necessarily reflect the views of the European Commission.

another oncology indication.²⁴ This methodology adopts an MCDA “clean-slate” approach for HTA¹² based on multi-attribute value theory (MAVT),^{28,29} with the aim to derive a comprehensive value function following five iterative phases: 1) problem structuring, 2) model building, 3) model assessment, 4) model appraisal, and 5) development of action plans.²³ The AVF was operationalized using the decision support system M-MACBETH, enabling the use of visual graphics to build a model of values, acting as a facilitation tool to inform both the design phases (1 and 2) and evaluation phases (3 and 4) of the methodological process.³⁰ Additional information on the methodological process can be found in the Online Supplementary Appendix.

Clinical Practice and Scope of the Exercise (Problem Structuring)

In 2004, clinical evidence demonstrated the survival benefit of docetaxel chemotherapy in combination with prednisolone for castrate-resistant prostate cancer (CRPC) patients, when compared with mitoxantrone in combination with prednisolone.^{31,32} Since then a fruitful period in the development of prostate cancer drugs followed with a number of new molecules indicating significant clinical benefit in patients previously treated with docetaxel including abiraterone, cabazitaxel, and enzalutamide.^{33–35} Cross-resistance appears to exist between abiraterone and enzalutamide, indicating that it is unlikely to experience clinical benefit by switching from one treatment to the other,^{36,37} based on which NICE recommends only the use of either treatment and not both.^{38,39} Although specific patient subpopulations might be contraindicated to any of the treatments, in most cases all three new and expensive drugs could be suitable for patients, and therefore, the topic of CRPC treatment coverage acts as a suitable decision context for the application of the value framework.

The analysis presented here focuses on assessing the value of third-line treatments for mCRPC following prior docetaxel-containing (i.e., second line) chemotherapy, essentially adopting the clinical practice guidelines from the European Society for Medical Oncology (ESMO) and the scopes of the respective Technology Appraisals (TAs) for each technology by NICE and TLV. In doing so, the same clinical evidence from the corresponding TAs was used to inform the respective criteria of the value tree, but this was further supplemented by additional evidence for value concerns not explicitly addressed by NICE and TLV. The scope of NICE TA255 was adopted for the case of cabazitaxel in

combination with prednisone, the scope of NICE TA259 and TLV TA4774/2014 was adopted for the case of abiraterone in combination with prednisone, whereas the scope of NICE TA316 and TLV TA2775/2013 was adopted for the case of enzalutamide.^{38–42}

Adaptation of the Advance Value Tree for Metastatic Prostate Cancer (Model Building)

Selection of evaluation criteria benefitted from a hybrid approach, which essentially combined features from “value focused thinking”⁴³ and “alternative focused thinking.”⁴⁴ Initially, a top-down approach was adopted for the selection of a general set of evaluation criteria by decomposing the overall value concerns of HTA decision makers into subconcerns, which took place prior to the selection of any drug options; this was followed by the adaptation of the general criteria into disease-specific criteria, which took place after comparing the characteristics of the particular drugs evaluated, as part of a bottom-up approach.²³

More precisely, following a systematic literature review and expert consultation as part of the Advance-HTA project,^{4,1} a generic value tree was developed outlining higher level criteria that were decomposed into lower level criteria for assessing the value of new medicines in the HTA context.²² The structure of this generic value tree, the Advance Value Tree, consists of five value domains or criteria clusters, aiming to reflect all the essential value attributes of new medicines in tandem with their indication under a prescriptive decision-aid approach. These include 1) burden of disease (BoD), 2) therapeutic impact (THE), 3) safety profile (SAF), 4) innovation level (INN), and 5) socioeconomic impact (SOC), with the first value domain relating to the disease of interest and the four latter domains relating to the actual medical intervention(s), essentially producing the following value function:

$$Value = f(BoD, THE, SAF, INN, SOC) \quad (1)$$

In consultation with a clinical oncologist, the generic value tree was then adapted for the case of mCRPC

i Advance-HTA was a research project funded by the European Commission’s Seventh Research Framework Programme (FP7). It was comprised of several complementary streams of research that aimed to advance and strengthen the methodological tools and practices relating to the application and implementation of HTA. It was a partnership of 13 consortium members led by the London School of Economics—LSE Health.

while striving to adhere to the ideal decision theory properties to ensure methodological robustness. As a result, a disease-specific value tree was produced with five criteria clusters decomposed into eight subcriteria clusters and a total of 18 emerging criteria attributes (Supplementary Figure A1 in the Online Appendix). The definitions of all criteria attributes are listed in Table 1.

Subsequently, the value tree was validated together with the decision conference participants as part of a more constructive decision-aid approach allowing for interactivity and enabling group learning. This “socio-technical” approach is reflecting another type of mixed approach adopted in the evaluation (i.e., values construction): initially, model adaptation and structuring took place in consultation with a clinical expert, which was then collectively validated and evaluated with the group before the completion of the model building phase.⁴⁵

Evidence Considered and Alternative Treatments Compared (Model Building)

The alternative treatment options compared in the analysis were abiraterone (Zytiga) in combination with prednisone, cabazitaxel (Jevtana) in combination with prednisone, and enzalutamide (Xtandi) monotherapy.

The evidence sources used included the peer review publications concerning the pivotal clinical trials of the alternative treatment options that were considered in the appraisals of NICE and TLV,^{33–35,46} the NICE Evidence Review Group (ERG) reports,^{47,48} or their relevant peer review publications,^{49,50} a Swedish population study on health related quality of life,⁵¹ the Product Information sections of the European Public Assessment Reports (EPAR) from the European Medicines Agency (Annex I and III),^{52–54} the Anatomical Therapeutic Chemical (ATC) classification system indexes available through the portal of the WHO Collaborating Centre for Drug Statistics Methodology,⁵⁵ the US National Library of Medicine database of clinical trials available through ClinicalTrials.gov,⁵⁶ and the British National Formulary (BNF).⁵⁷ The clinical evidence (falling under the THE and SAF criteria clusters) for each alternative treatment option were sourced from the same clinical studies that NICE and TLV evaluated (single pivotal trials for each drug). The source of evidence used for identifying the performance of options across the evaluation criteria is shown in Table 1. Additional information on the evidence considered can be found in the Online Appendix.

Setting Attribute Ranges and References Levels (Model Building)

With regard to model building, the allowable range of performance across the different attributes (i.e., attribute ranges) was selected based on the performance of the alternative treatment options compared. Within the attribute ranges, “lower” (x_l) and “higher” (x_h) reference levels were defined to serve as benchmarks for preference value scores of 0 and 100, respectively, acting as value anchors for constructing criteria value functions and eliciting their relative weights.^{58,59,ii} The “lower” reference levels denoted less preferred levels aiming to reflect “satisfactory” performance, whereas the “higher” reference levels denoted more preferred levels aiming to reflect “ideal” performance. The latter was defined by considering plausible changes in best available performance, for example, by accounting for new product characteristics that might enter the market in the near future that could outperform the current ones. The reference levels of the clinical attributes (under the THE and SAF value domains) were decided in consultation with a clinical expert (urologic oncologist). Depending on the criterion under consideration, the typical rationale adopted the BSC performance as a “satisfactory” reference level with a hypothetical performance improvement of 20% better than the best available option acting as the “ideal” reference level (or vice versa for negative attributes). The 20% hypothetical improvement was selected because it was deemed to be a realistically plausible clinical effect of forthcoming products. By considering the performance of best available option(s) among the treatments evaluated in the exercise and accounting for plausible performance improvement in the near future, the value scale essentially reflected characteristics of both a “local” scale (benchmarks derived based on the performance of options evaluated, i.e., *what is best available*) and a “global” scale (benchmarks derived based on the plausible performance of options not necessarily captured in the exercise, i.e., *what is best plausible*).⁶⁰ The emerging value scores of options could in theory take negative or higher than 100 values where $v(x_{\text{lower}}) = 0$ and $v(x_{\text{higher}}) = 100$, essentially by conducting a linear transformation that is acceptable as an interval scale such as a value scale. “Lower” and “higher” reference levels for all attributes at the pre-workshop stage and the basis of their selection are outlined in Table A1 in the Online Appendix (assuming

ii These are interval scales and therefore it is important to set up clear bounds/limits for each attribute.

Table 1 Attributes Definition and Sources of Evidence

Cluster	Attribute	Definition	Evidence Source		
			Abiraterone	Cabazitaxel	Enzalutamide
Therapeutic Impact	Overall survival	The median time from treatment randomization to death	de Bono et al. (2011)	de Bono et al. (2010)	Scher et al. (2012)
	Health-related quality of life	Health-related quality of life using the EQ-5D score	TA259; TA316	N/A—assumed	TA316
Safety Profile	Radiographic tumor progression	The median survival time during which patients have not experienced disease progression (using RECIST criteria)	de Bono et al. (2011)	de Bono et al. (2010)	Scher et al. (2012)
	PSA response	The proportion of patients having a $\geq 50\%$ reduction in PSA	Fizazi et al. (2012)	de Bono et al. (2010)	Scher et al. (2013)
Innovation Level	Treatment discontinuation	The proportion of patients discontinuing treatment due to AEs	de Bono et al. (2011)	de Bono et al. (2010)	Scher et al. (2012)
	Contraindications	The existence of any type of contraindication accompanying the treatment	EPAR, Prescribing info	EPAR, Prescribing info	EPAR, Prescribing info
Socioeconomic Impact	ATC Level 1	The technology's relative market entrance with regard to its ATC Level 1 (Anatomical)	WHO ATC index	WHO ATC index	WHO ATC index
	ATC Level 2	The technology's relative market entrance with regard to its ATC Level 2 (Therapeutic)	WHO ATC index	WHO ATC index	WHO ATC index
	ATC Level 3	The technology's relative market entrance with regard to its ATC Level 3 (Pharmacological)	WHO ATC index	WHO ATC index	WHO ATC index
	ATC Level 4	The technology's relative market entrance with regard to its ATC Level 4 (Chemical)	WHO ATC index	WHO ATC index	WHO ATC index
	ATC Level 5	The technology's relative market entrance with regard to its ATC Level 5 (Molecular)	WHO ATC index	WHO ATC index	WHO ATC index
Phase 1	Phase 1	The number of new indications for which the technology is investigated in Phase 1 clinical trials	ClinicalTrials.gov	ClinicalTrials.gov	ClinicalTrials.gov
	Phase 2	The number of new indications for which the technology is investigated in Phase 2 clinical trials	ClinicalTrials.gov	ClinicalTrials.gov	ClinicalTrials.gov
	Phase 3	The number of new indications for which the technology is investigated in Phase 3 clinical trials	ClinicalTrials.gov	ClinicalTrials.gov	ClinicalTrials.gov
Marketing authorization	Marketing authorization	The number of new indications that the technology has gained an approval for at the stage of marketing authorization	ClinicalTrials.gov	ClinicalTrials.gov	ClinicalTrials.gov
	Delivery posology	The combination of the delivery system (route of administration and dosage form) with the posology (frequency of dosing and duration of administration) of the treatment	EPAR, Prescribing info	EPAR, Prescribing info	EPAR, Prescribing info
Special instructions	Special instructions	The existence of any special instructions accompanying the administration of the treatment	EPAR, Prescribing info	EPAR, Prescribing info	EPAR, Prescribing info
	Medical costs impact	The impact of the technology on direct medical costs excluding the purchasing costs of the technology	BNF 69, Prescribing info, Connock et al. (2011), Riemsa et al. (2013), TA259	BNF 69, Prescribing info, de Bono et al. (2010), TA255	BNF 69, TA316

AE, adverse event; ATC, Anatomical Therapeutic Chemical; EPAR, European Public Assessment Reports; PSA, prostate-specific antigen; WHO, World Health Organization.

no impact of LHRH analogue). Additional information on the methodological basis adopted for setting the min-max attribute limits and reference levels is provided in the Online Appendix.

Decision Conference (Model Assessment and Appraisal)

The model assessment and model appraisal phases of the exercise took place through a facilitated workshop with assessors from TLV, taking place as a decision conference and hosted at the head offices of TLV in Stockholm, Sweden. In total four experts participated, including one medical investigator, one pharmacist, and two health economists. Background material introducing the scope of the exercise in more detail was sent to the participants one week before the workshop.

The author acted as an impartial facilitator and assisted the group's interaction and thinking about the decision problem using the preliminary version of the mCRPC specific value tree (Supplementary Figure A1 in the Online Appendix) and the relevant data as the model's starting point, based on which value judgements and preferences were elicited on-the-spot.^{27,61,62} On the day of the workshop the preliminary model was validated with the participants by revising it cluster by cluster in real time through an open discussion, seeking group consensus by adopting an iterative and interactive-model-building process where debate was encouraged and differences of opinion actively sought. More information on the decision conference is provided in the Online Appendix.

MCDA Technique (Model Assessment and Appraisal)

The AVF adopts a value measurement MCDA methodology making use of a simple additive (i.e., linear, weighted average) aggregation model that assumes preference independence between the different attributes, with overall value $V(\cdot)$ of an option a defined by the equation below²²:

$$V(a) = \sum_{i=1}^m w_i v_i(a) \quad (2)$$

where m is the number of evaluation criteria, $w_i v_i(a)$ is the weighted partial value function of evaluation criterion i for option a , and $V(\cdot)$ is a multi-attribute value function based on value theory principles.²⁹

Each attribute's value function correlates the performance along an attribute range to a value scale. These were elicited from the workshop participants using the Measuring Attractiveness by a Categorical Based Evaluation Technique (MACBETH), based on MAVT.²⁸ This is a pairwise qualitative comparison approach where qualitative judgements about the difference of value between different pairs of attribute levels (e.g., difference in value between 3 and 6 months overall survival) are expressed using seven categories (no difference, very weak difference, weak difference, moderate difference, strong difference, very strong difference, or extreme difference in value).^{63,64} By only requiring such semantic judgements, which are then converted into a cardinal (interval) value scale using a mathematical algorithm with a set of decision rules, MACBETH provides a simple, interactive, and constructive approach, with various real-world applications illustrating its usefulness as a decision support tool⁶⁵⁻⁶⁸; the mathematical foundations of MACBETH have been extensively described.⁶⁹ A cardinal scale can reveal information about the difference between its levels, in contrast to an ordinal scale that can only be used for a more restrictive ranking along its levels without allowing any arithmetic operations. An indirect (qualitative) swing weighting technique was applied involving the MACBETH value judgements to elicit relative criteria weights,⁶⁴ given that direct questions of importance for a criterion that do not take into account their attribute ranges are known to be one of the most common mistakes in making value trade-offs.⁷⁰

In turn, the M-MACBETH software was used as a support system in order to construct the value tree, elicit criteria value functions (through which options were scored), assign relative criteria weights through the MACBETH qualitative swing weighting protocol, combine preference value scores and weights together using additive aggregation to derive overall weighted preference value (WPV) scores, and perform sensitivity analysis on criteria weights.^{58,71} The software automatically performs consistency checks between the qualitative judgements expressed, and in addition a second consistency check was manually performed by the author to validate the cardinality, i.e. the interval nature, of the emerging value scale. This was done by comparing the sizes of the intervals between the proposed scores and inviting participants to adjust them if necessary,⁷² a requirement that is essential for the application of simple additive value models. More information regarding the technical details of MACBETH is available in the Online Appendix.

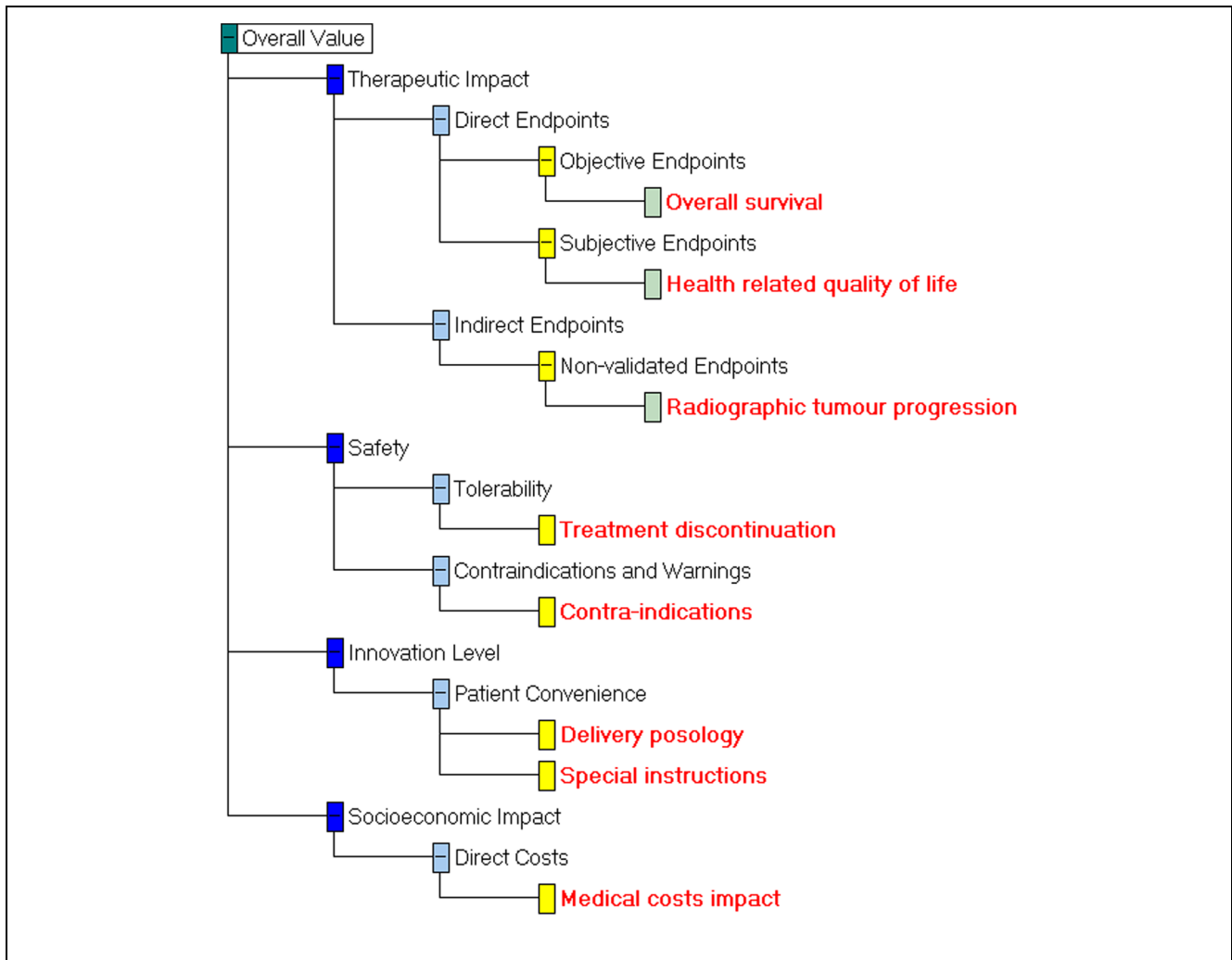


Figure 1 Final value tree for metastatic prostate cancer (post-workshop). Image produced using the M-MACBETH (beta) software, version 3.0.0.

Costs Calculation (Model Assessment and Appraisal)

Drug costs were calculated according to UK prices (excluding VAT), pack sizes, and dosage strengths as found in the BNF⁵⁷; the recommended dosages and treatment durations were taken from the pivotal trial for the case of cabazitaxel,³³ the respective NICE technology appraisal for the case of enzalutamide,³⁹ and the labelling and package leaflet (EPAR—Annex III) for the case of abiraterone.⁵⁴ Vial wastage was assumed in all calculations. Drug administration costs for cabazitaxel were kept consistent with the respective NICE TA.⁴⁰ The UK perspective was adopted as a neutral benchmark, partially because of the readily available data and primarily in order to allow the measurement

of cost(s) in a common unit across a series of similar case studies in different European countries, so that overall WPV scores can then be viewed against the same cost denominator to produce comparable cost-value ratios.

Results

Criteria Validation and Amended Value Tree for Metastatic Prostate Cancer

Overall, consensus was reached in terms of criteria inclusion and exclusion with no criteria considered to be missing. The final version of the value tree, as emerged following the open discussion with the participants of the workshop, is shown in Figure 1. In total,

Table 2 Options Performance Across the Criteria Attributes

Attribute Name	Attribute Metric	Lower Level	Abiraterone	Cabazitaxel	Enzalutamide	Higher Level
Overall survival	Months	13.6	15.8	15.1	18.4	22.1
Health-related quality of life	Utility (EQ-5D)	0.72	0.76	0.76 ^a	0.76	0.82
Radiographic tumor progression	Months	2.9	5.6	8.8	8.3	10.6
Treatment discontinuation	% of patients	10	19	18	8	0
Contraindications	Types of contraindications	Hyp + hep imp + low neut	Hyp + hep imp	Hyp + hep imp + low neut	Hyp	None
Delivery posology	Types of delivery system and posology combinations	Oral, daily—one off + IV, every 3 weeks—1 hour	Oral, daily—one off	Oral, daily—one off + IV, every 3 weeks—1 hour	Oral, daily—one off	Oral, daily—one off
Special instructions	Types of special instructions	Concomitant and/or pre-med + no food	Concomitant and/or pre-med + no food	Concomitant and/or pre-med	None	None
Medical costs impact	GBP (£)	10,000	5,750	7,992	567	0

Hep imp, hepatic impairment; hyp, hypersensitivity; low neut, low neutrophil count.

^aUsed the same score of the other two options as data not available.

10 out of the 18 attributes were removed from the value tree because they were judged from the participants to be nonfundamental for the scope of the exercise. Most of these attributes were part of the innovation level cluster, with all attributes relating to the ATC classification system (five attributes) and spillover effects (four attributes) being excluded. The only other attribute removed was the PSA (prostate-specific antigen) response attribute of the therapeutic impact cluster because it was thought to be redundant, without adding any value on top of the remaining therapeutic attributes.

An example of a MACBETH value judgements matrix, its conversion into a value scale, and the respective scoring of the alternative treatments for the case of Overall Survival (OS) criterion is described in the Online Appendix. Effectively all attributes produced almost linear value scales, either based on quantitative or qualitative performance levels.

Options Performance, Criteria Weights, and Overall Preference Value Rankings

The performance of the options across the different attributes together with the “lower” and “higher” reference levels are shown in Table 2.

The overall WPV scores of the options and their breakdown into their partial value scores across the different criteria attributes with their respective weights are shown in Table 3. Enzalutamide scored the highest overall WPV score of 58.7. Abiraterone and cabazitaxel produced overall WPV scores of 6.9 and 1.4, respectively, partially due to their “negative” performance in the treatment discontinuation attribute, producing absolute preference value scores of -95.3 and -87.5 , respectively, and WPV scores of -20.2 and -18.6 , respectively. This was due to their performance on the treatment discontinuation attribute (19% and 18%, respectively), which lay below the lower reference level of the scale (i.e., 10%). A stacked bar plot of the WPV scores of the alternative treatments across the attributes is shown in Figure 2.

The relative weights of importance assigned to the different attributes are shown in Figure 3. By taking into account the “lower” to “higher” ranges of the attributes, the greatest relative weights were yielded for the case of Overall Survival, Treatment Discontinuation, and Health-Related Quality of Life (with relative magnitudes of 23, 21, and 15 out of 100, respectively), adding up a combined relative weight of 60% of the total. With regard to the total weights assigned across the criteria

Table 3 Overall Weighted Preference Value (WPV) Scores, Partial Preference Value Scores, Relative Weights, Costs, and Cost per Unit of Value

	Lower Level	Abiraterone	Cabazitaxel	Enzalutamide	Higher Level	Relative Weights
Overall WPV score	0	6.9	1.4	58.7	100	100
Overall survival	0	26.2	17.9	56.3	100	23
Health-related quality of life	0	40.0	40.0	40.0	100	15
Radiographic tumor progression	0	26.5	74.0	66.8	100	6
Treatment discontinuation	0	-95.3	-87.5	20.0	100	21
Contraindications	0	33.3	0.0	83.3	100	12
Delivery posology	0	100.0	0.0	100.0	100	3
Special instructions	0	0.0	60.0	100.0	100	4
Medical costs impact	0	40.5	19.1	93.7	100	15
Costs (£)		21,900	23,900	24,600		
Cost per unit of value		3,173	17,509	419		

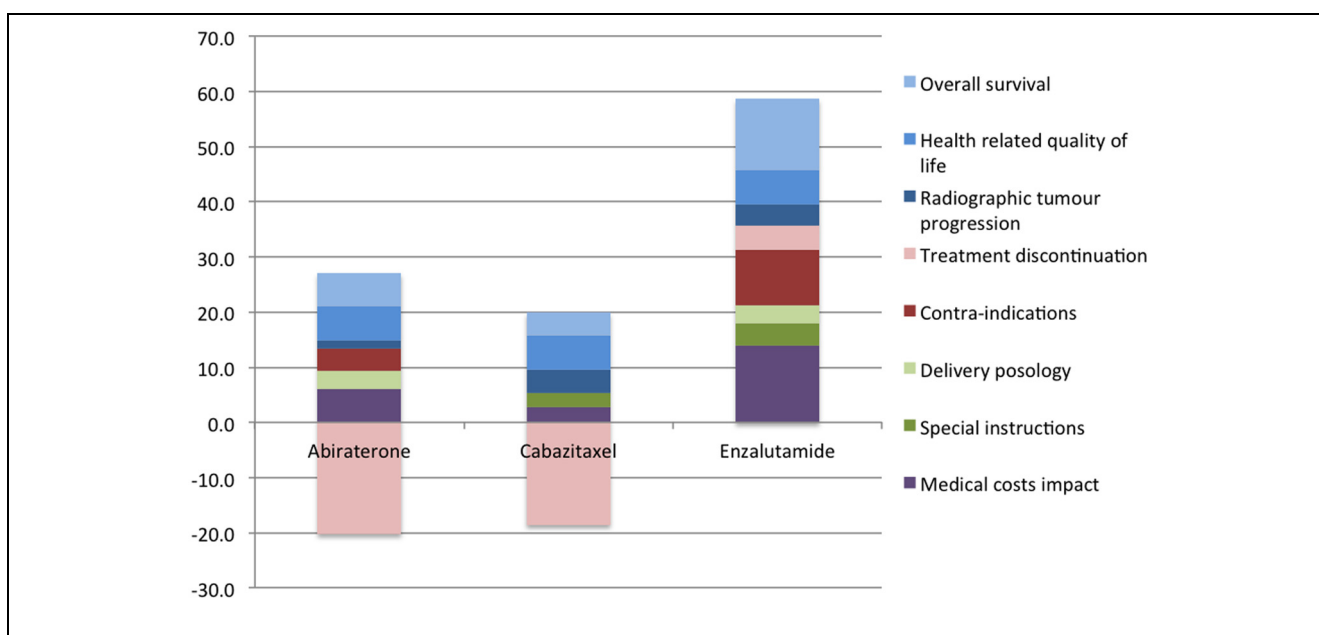


Figure 2 Stacked bar plot of treatments’ weighted preference value (WPV) scores across all attributes.

clusters, Therapeutic Impact cluster totaled a relative weight of 44 (three attributes), the Safety Profile cluster totaled a relative weight of 33 (two attributes), the Socioeconomic Impact cluster totaled a relative weight of 15 (single attribute), and the Innovation Level cluster totaled a relative weight of 7 (two attributes) out of 100.

Value for Money Analysis

By using rounded up cost figures of £24,600 for enzalutamide, £21,900 for abiraterone, and £23,900 for cabazitaxel (£22,190 drug cost and £1,710 administration cost)

and dividing them with overall WPV scores, the costs per MCDA value unit were calculated to be £419, £3,173, and £17,509, respectively (Table 3). Therefore, in terms of value-for-money, cabazitaxel is shown to be dominated by abiraterone, being also very close to be dominated by enzalutamide (£500 difference based on the prices used). Enzalutamide on the other hand is associated with a higher cost (£2,500 difference) and a higher overall WPV score (51.8 difference) compared with abiraterone. The overall WPV scores of the alternative treatments versus their costs (purchasing costs plus any administration costs) is provided in the form of a cost-benefit plot in Figure 4.

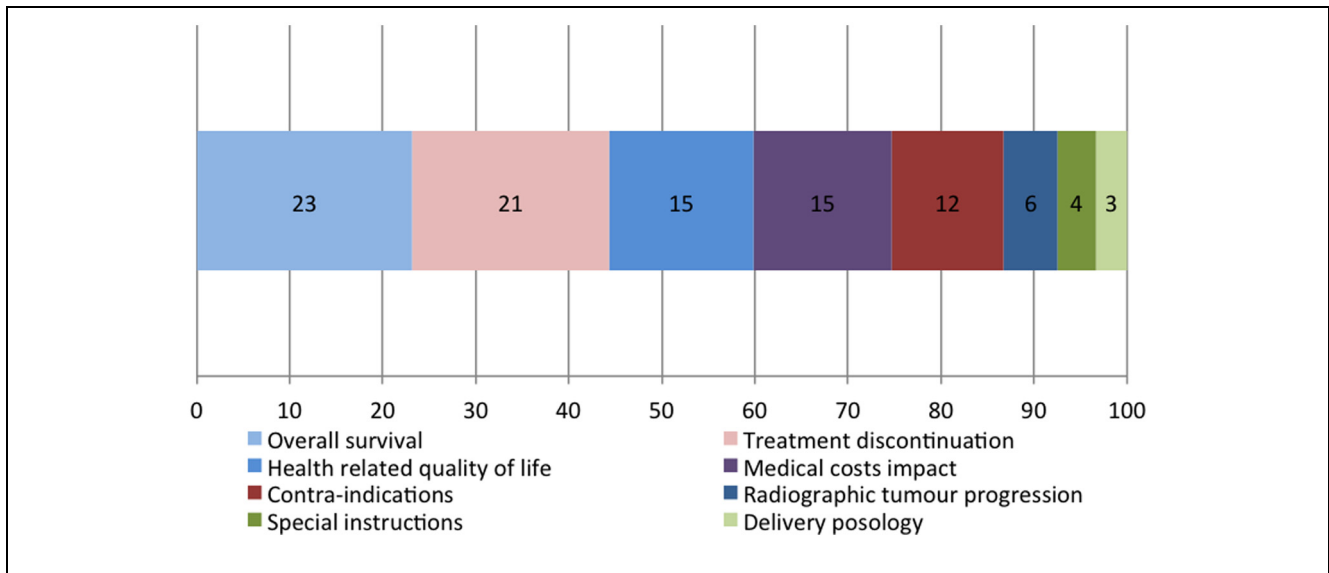


Figure 3 Relative criteria weights stacked bar.

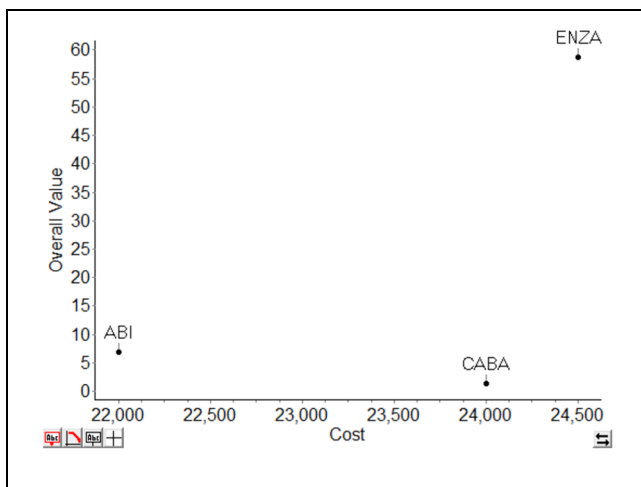


Figure 4 Cost-benefit plot of overall weighted preference value (WPV) scores versus costs. Image produced using the M-MACBETH (beta) software, version 3.0.0.

Sensitivity and Robustness Analysis

Parameter uncertainty relating to the estimation of weights was addressed by conducting deterministic sensitivity analysis. At the end of the workshop, the impact of baseline weights' changes on options' rankings was explored using one-way sensitivity analysis. In order for cabazitaxel to become better ranked than abiraterone, PFS relative weight would have to change from 5.8 to

15.6, Treatment Discontinuation relative weight from 21.2 to 53.9, or Special Instructions relative weight from 4.1 to 12.2. Alternatively, for cabazitaxel to become better ranked than enzalutamide, PFS relative weight would have to change from 5.8 to more than 89.5. Under no scenario abiraterone could have been ranked above enzalutamide. Therefore, conclusions were robust as the ranking of the treatments was not sensitive to single variations of up to at least 100% along the attributes weight range. The robustness of the results was also validated by conducting seven-way sensitivity analysis in the reference levels of the quantitative attributes using the respective function of the M-MACBETH software ("Robustness analysis"). A simultaneous change of up to $\pm 10\%$ across all of the attribute reference levels would not affect the ranking of the alternative treatments. This means a change of 20 points in the value scale between the highest and the lowest scores for every attribute (i.e. the allowable range of eligible scores), did not affect treatments' ranking. In terms of weighting, a simultaneous change of $\pm 3\%$ would be needed for a change in the rankings of the options to be observed (cabazitaxel to become second, abiraterone to become third), with no changes in the best-ranked option plausible.

Other types of uncertainty also exist, including stochastic uncertainty, structural uncertainty, and heterogeneity that could possibly be addressed with other approaches (or combinations of), such as probabilistic sensitivity analyses, Bayesian frameworks, fuzzy set theory, or grey theory.⁷³ For instance, if significant

uncertainty exists with regard to options' performance due to sampling variation from clinical trials, or in terms of criteria weights due to failure to agree on them, the application of point estimates might be inappropriate in which case stochastic multi-criteria acceptability analysis (SMAA) could be used.⁷⁴

Discussion

This case study demonstrates the application of a new value framework, the AVF,²² adopting an MCDA methodological process in practice²³ for a set of alternative mCRPC treatments from the perspective of an HTA agency. Generally, MCDA methods in the context of HTA could be applied using two different ways, either as part of a supplementary/incremental approach or a pure/clean-slate approach.^{12,23} The former could in theory be used in combination with economic evaluation techniques to adjust the incremental cost effectiveness ratio (ICER) through the incorporation of additional parameters of value (or benefit criteria), therefore essentially "tweaking" the QALY or any other relevant outcome measure. The latter refers to deriving value anew, not necessarily involving the QALY or other value metric from some type of economic evaluation. The current application of AVF as part of this case study focuses on the latter.

In terms of design, implementation, and review of the analysis, the process adopted is effectively in alignment with the recent ISPOR good practice guidelines on the use of MCDA for health care decisions.⁷⁵ Through a decision conference, alternative options were assessed and ranked based on their overall WPV scores reflecting their performance against a set of evaluation criteria weighted for their relative importance based on the preferences of the group, yielding a holistic benefit component. In doing so, MACBETH was used as an MCDA modelling technique to operationalize the AVF in terms of scoring the treatment options and weighting the relative importance of the evaluation criteria. Subsequent consideration of drug costs (purchasing and administration costs) enabled the estimation of "cost per unit of value" ratios, which showed a dominance between the second and third ranked options. Assuming the existence of well-defined budgets, this approach could be used to derive MCDA-based cost-benefit ratios, prioritizing the most efficient funding options with the lowest cost-benefit ratios until the available budget is exhausted.⁷⁶

The TLV assessors who participated in the workshop gave positive feedback about the potential usefulness of the value framework, having the prospects of acting as a

decision support tool in the evaluation of new medicines. Main advantages included the feasibility to transparently assess the performance of the options across a number of explicit evaluation criteria, while allowing to elicit their value trade-offs (i.e., their relative importance), but also its overall facilitative nature in the construction and analysis of the groups' value preferences.

It should be highlighted that the decision context addressed in this case study is a one-off evaluation problem within a particular disease indication in contrast to the typical remit of TLV, which involves repeated decisions around the reimbursement of drugs across different disease areas by considering their opportunity cost through the use of, flexible or adjusting, willingness to pay thresholds. Consequently, even the most valuable or efficient treatment in the exercise could prove to be non-cost-effective on traditional economic evaluation grounds. This notion would be supported by the fact that the breath of the value model captured aspects that go beyond the formal remit of TLV. Therefore, the results aim to illustrate the methodology's feasibility rather than to inform policy making.

Even under the more restricted context of feasibility testing, the findings should be interpreted under the specified MCDA characteristics of the exercise, namely, the particular decision context defined and the value model developed. The latter consists of the alternative options assessed, the evaluation criteria incorporated, the relevant value scales defined, the respective value functions elicited, and the relative weights elicited. Although more challenging, the value framework could also be applied at broader inter-indication levels, aiming to assess treatments for different indications that might require the use of more generic clinical evaluation criteria and less disease-specific endpoints, as, for example, the use of QALY that allows the comparison of health gain across different populations.

With regard to the removal of the spillover effect attributes under the innovation level cluster, this took place on the grounds that they "currently go beyond the agency's remit" and therefore should not be considered in the first place. However, the opinion was evident that in real practice this information could act as supplementary type of evidence, not for influencing the coverage of the respective treatments but possibly for other secondary purposes as, for example, awarding reimbursement extensions or communicating internally possible new indications in the near future.

The following section discusses some limitations of the case study and focuses on challenges faced by MCDA applications in HTA settings and how they could be

encountered. A more extensive discussion around the types of strengths and opportunities possibly offered by the AVF are provided elsewhere, both in terms of relevant practice and policy implications as part of a recent case study on another oncology indication²⁴ and from a more theoretical perspective.²²

Limitations of the Case Study

The results from the case study should be interpreted with caution as they comprise the outcome of a simulation exercise aiming to test the AVF in practice by illustrating its application rather than to inform real coverage decision making in respect to the particular drugs under consideration. Given the absence of head-to-head clinical trials comparing directly all treatments of interest and the lack of relative treatment effects across the clinical attributes of interest, un-synthesized evidence from the respective single pivotal trials of the alternative treatments were used. However, using un-synthesized evidence from different clinical trials to directly compare alternative treatments is not accurate, even if the populations of patients across the different trials are similar (in terms of disease severity and treatment history). Ideally, an indirect treatment comparison should be conducted first by using a common comparator to estimate the relative effects of two treatments versus the comparator, or a network meta-analysis that combines both direct and indirect evidence available through a mixed treatment comparison.⁷⁷ Therefore, an important limitation of this study is the lack of relative effects as part of clinical evidence and the use of their absolute effects from different clinical trials with the assumption that they can be directly compared. In real-world evaluations, evidence synthesis would be required to take place together with evidence collection as part of the model building phase. An example can be illustrated through the application of an SMAA approach for assessing the comparative benefit-risk of alternative statins in primary prevention,¹⁹ using comparative effects from evidence of three meta-analyses,^{78–80} or the combination of SMAA with a network meta-analysis for assessing the comparative benefit-risk of second-generation antidepressants.⁸¹

The relative small number of participants could be perceived as another limitation on the grounds of insignificant representation of perspectives for the purpose of informing policy making, but capturing an all-round set of preferences was not among the primary aims of the exercise. Ideally, a group size of between 7 and 15 participants should be aimed for; these sizes allow efficient group processes to emerge while preserving individuality, as they are large enough to

represent all major perspectives but small enough for working toward agreement.²⁶ However, it should be noted that in reality final TLV decisions are made by a separate expert board/committee within the agency (the Pharmaceuticals Benefit Board), which consists of only seven members and whose overall expertise is well aligned with the expertise of the exercise participants.⁸²

Another issue was that only the HRQoL of the stable disease state was assessed because none of the treatments was assumed to have any effects during the progressive disease state.³⁹ In other disease indications this might not hold true in which case the relevant HRQoL attribute would need to capture both the stable and progressive disease states, possibly in the form of an OS-HRQoL aggregated attribute (given their possible preference dependency, discussed below), therefore producing a similar metric to QALYs.

Furthermore, another issue which could be perceived as a limitation was the definition of the “lower” reference level in the treatment discontinuation attribute and its impact on scoring, as it could influence the negative partial value scores observed in two of the treatments and consequently their overall WPV scores. The lower reference level of 10% used on the basis of Best Supportive Care performance was sourced from the placebo comparator arm of enzalutamide’s pivotal clinical trial (*AFFIRM*). The choice of using the placebo arm of the *AFFIRM* trial to proxy BSC performance and not the comparator arm from any of the other two treatments’ pivotal trials was because it better resembled BSC; in abiraterone’s pivotal trial (*COU-AA-301*) all patients in the placebo comparator arm were also administered steroids (prednisone), whereas in cabazitaxel’s pivotal trial (*TROPIC*) all patients in the comparator arm were administered steroids (prednisone or prednisolone) on top of chemotherapy (mitoxantrone). As a result, abiraterone and cabazitaxel produced negative partial value scores in treatment discontinuation as their performance lay below the lower reference level. The basis adopted for setting all clinical attribute reference levels is extensively described in the Methods section, and although it took place in consultation with a clinical expert, others might have chosen different attribute levels. However, any such differences would most probably have been of minor impact, not necessarily affecting the overall valuation or ranking of the treatments.

Challenges of MCDA Applications in HTA

Among the main challenges of MCDA applications in HTA evaluations include the technical requirements in

ensuring that all attributes possess the necessary theoretical properties.^{83,84}

Because of the use of a simple additive (i.e., weighted average aggregation) model, the violation of preference-independence is of particular relevance as it might undermine the validity of such models.⁸³ There are two types of preferential independence, at ordinal and cardinal levels, the former of which is a precondition for the latter, but not vice versa. Intuitively, ordinal preference independence means that alternative options can be ranked on one dimension (e.g., attribute) independently of their impact on another dimension. In contrast, the notion of cardinal preference independence, which is also referred to as “difference independence,”⁸⁵ denotes that the difference in attractiveness—or added value of an improvement—between two performance levels on one criterion (e.g., attributes levels) does not depend and can be measured independently of the levels on another criterion.⁸⁶ Such a “difference consistency” between (the values of the) attributes is often taken as a working hypothesis given that it is “so intuitively appealing that it could simply be assumed to hold in most practical applications.”^{87(p284)}

However, evidence suggests potential preference dependencies between health gain and disease severity,⁸⁸ or overall survival and HRQoL.²² Considering the absence of a disease severity criterion, in order to avoid potential preference-dependence between the OS and HRQOL attributes and given an equal performance across all options in terms of the latter, OS value functions derived were effectively conditional on the range of the HRQoL attribute at a cardinal level. More precisely, the “lower” (x_l) and “higher” (x_h) reference levels of the plausibly dependent HRQoL attribute (EQ-5D index scores) were revealed to participants, that is, $v(x_{-0.72}) = 0$ and $v(x_{-0.82}) = 100$, and they were instructed to assume an identical performance between the options on the dependent attribute (given that all three options had identical performance), but without disclosing the exact figure. However, it should be noted that the use of such “conditionally cardinal value functions” might in theory be incompatible with the use of an additive model and therefore should only be used with caution, as, for example, within descriptors of performance (i.e., attributes).⁸⁹ Alternatively, a multiplicative model could be used; see the example of the Thai essential list.⁹⁰ In any case, the use of a simple test involving the combination of Reasoning Maps with MACBETH that takes the form of a multilinear evaluation model structure has been proposed as a way to capture potential value interactions between criteria and therefore cardinal preference dependencies.⁹¹

Other technical challenges include a possible lack of connection between attribute ranges (i.e., criteria scales) and weights, and the incorporation of economic evaluation-related considerations in the MCDA.⁸⁴ First, assigning weights of “generic importance” without considering the performance of the alternative decision options being analyzed is known to be one of the most common mistakes when making value trade-offs.^{43,70} In the value framework adopted, elicitation of weights takes place following the selection of the alternative treatments and the specification of their performance as part of the adaptation of the generic Advance Value Tree into a disease-specific value tree.²² Second, considering the costs and/or the cost-effectiveness of the options can lead to theoretical challenges as costs do not represent attributes of benefit whereas the effectiveness component would have already been captured as part of the relevant criteria, therefore possibly leading to double-counting.⁹² Combining both benefits and costs together can also cause practical hurdles when trying to link the MCDA results with policy recommendations, as prioritizing resource allocation would not be a clear task given the absence of a cost-benefit (or value) ratio that reflects efficiency. Instead, the value framework applied only considers impact on resources without accounting for the purchasing cost of the options themselves,²² therefore providing a more comprehensive unit of benefit. Such benefit units could then be considered in alignment with the respective costs of the options to derive appropriate cost-value ratios, which could be used to inform the allocation of resources by understanding the opportunity cost of different decisions assuming a defined budget,⁷⁶ possibly as part of a portfolio decision analysis exercise.

Future research could investigate the opportunity cost associated with disinvestments which could feed the development of incremental cost value ratio thresholds, acting as potential cutoff points for efficiency similarly to current ICER thresholds used in economic evaluations. However, this would face all the current challenges associated with the technical difficulties in estimating sound ICER thresholds to reflect opportunity cost.^{93,94}

Conclusion

In health care systems facing significant budgetary pressures, HTA challenges relating to the evaluation of and resource allocation for innovative treatments require novel value-based assessment methodologies that can facilitate preference elicitation and decision making in a transparent way. Such methodologies should be based on robust theoretical principles so the results can lead to

credible decisions. Future research could help validate the robustness and usefulness of the current value framework by conducting similar case studies with other HTA and insurer bodies across different European countries.

In order to catalyze the possible implementation of any such MCDA-based value framework, even at a testing phase, the development of the model and the application of the methodology altogether should be aligned with the formal remit of the decision-makers' institutions. Establishing clear goals for the role of the results would be beneficial for the initial exploration of the methodology. In settings where strict guidelines and requirements exist around the application of specific HTA methods, as in Sweden with the use of economic evaluation, the exploration of a supplementary or "incremental" MCDA mode to cost-effectiveness analysis might act as a more realistic option for decision makers so that the results can be more readily adaptable to their current needs and therefore more easily implementable in daily practice. In turn, adaptation of a pure or "clean-slate" mode could still be viable but involving the utilization of more concise value trees that restrict the number of criteria around the formal value concerns currently in explicit use. This would allow capturing the primary value concerns as specified by the evaluation guidelines requirements while still offering the added benefits of assessing value through a decomposition approach that could enable the facilitation of decision making through a stepwise elicitation and transparent construction of value preferences. Following such a testing phase, decision makers could acquire important experience and confidence that could enable them to explore incorporating any additional value concerns judged to be relevant to their decisions.

Acknowledgments


The author is grateful to Prof Carlos Bana e Costa (University of Lisbon) for critically appraising this work and recommending ways of improvement. Constructive feedback and suggestions were also received from Dr. Monica Oliveira (University of Lisbon), Dr. Matteo Galizzi (London School of Economics), Prof. Larry Phillips (London School of Economics), Dr. Huseyin Naci (London School of Economics), Dr. Themistocles Resvanis (Massachusetts Institute of Technology), and Prof. Gilberto Montibeller (University of Loughborough), all of which are much appreciated. Special thanks go out to Dr. Mark Linch (University College London Hospitals) for providing valuable clinical insights that helped the author to understand the current clinical practice in England and Europe, informing the preliminary version of the model and earlier draft version of the respective section. The author is thankful to the staff of the Dental and Pharmaceutical Benefits Agency in Sweden (TLV)

that participated in two half-day workshops in Stockholm and whose preferences were elicited acting as the basis of the analysis: Loudin Daoura (medical investigator), Niklas Hedberg (chief pharmacist), Daniel Högberg (health economist), Douglas Lundin (chief economist). Special thanks go to Karl Arnberg (head of unit) and Douglas Lundin for helping organize the workshop. Olina Efthymiadou and Erica Visintin (London School of Economics) provided valuable research assistance with drug costs and innovation spillover effects calculations. The author finally thanks the participants at the PPRI conference of the WHO Collaborating Centre for Pharmaceutical Pricing and Reimbursement Policies in Vienna (October 2015) and the HTAi conference in Oslo (June 2015) for useful feedback. The vision of this research was originally proposed by Dr. Panos Kanavos (London School of Economics) as part of the Advance-HTA project. All outstanding errors are of the author.

Supplemental Material

Supplementary material for this article is available on the *Medical Decision Making* Web site at <http://journals.sagepub.com/home/mpp>.

ORCID iD

Aris Angelis  <https://orcid.org/0000-0002-0261-4634>

References

1. Drummond M, Tarricone R, Torbica A. Assessing the added value of health technologies: reconciling different perspectives. *Value Health*. 2013;16(1 suppl):S7–S13.
2. Devlin NJ, Lorgelly PK. QALYs as a measure of value in cancer. *J Cancer Policy*. 2017;11:19–25.
3. Wouters OJ, Naci H, Samani NJ. QALYs in cost-effectiveness analysis: an overview for cardiologists. *Heart*. 2015;101(23):1868–73.
4. Angelis A, Lange A, Kanavos P. Using health technology assessment to assess the value of new medicines: results of a systematic review and expert consultation across eight European countries. *Eur J Health Econ*. 2018;19(1):123–52. doi:10.1007/s10198-017-0871-0
5. Anderson JL, Heidenreich PA, Barnett PG, et al. ACC/AHA statement on cost/value methodology in clinical practice guidelines and performance measures: a report of the American College of Cardiology/American Heart Association Task Force on Performance Measures and Task Force on Practice Guidelines. *J Am Coll Cardiol*. 2014;63(21):2304–22.
6. Cherny NI, Sullivan R, Dafni U, et al. A standardised, generic, validated approach to stratify the magnitude of clinical benefit that can be anticipated from anti-cancer therapies: the European Society for Medical Oncology Magnitude of Clinical Benefit Scale (ESMO-MCBS). *Ann Oncol*. 2015;26(8):1547–73.

7. Schnipper LE, Davidson NE, Wollins DS, et al.; American Society of Clinical Oncology. American Society of Clinical Oncology statement: a conceptual framework to assess the value of cancer treatment options. *J Clin Oncol*. 2015; 33(23):2563–77.
8. Bach P. DrugAbacus App. Available from: <http://app.drugabacus.org/abacus-mskcc>
9. Malone DC, Berg NS, Claxton K, et al. International Society for Pharmacoeconomics and Outcomes Research comments on the American Society of Clinical Oncology Value Framework. *J Clin Oncol*. 2016;34(24):2936–7.
10. Angelis A, Kanavos P. Critique of the American Society of Clinical Oncology Value Assessment Framework for Cancer Treatments: putting methodologic robustness first. *J Clin Oncol*. 2016;34:2935–6.
11. Goetghebeur MM, Wagner M, Khoury H, Levitt RJ, Erickson LJ, Rindress D. Evidence and value: impact on DEcisionMaking—the EVIDEM framework and potential applications. *BMC Health Serv Res*. 2008;8:270.
12. Devlin N, Sussex J. *Incorporating Multiple Criteria in HTA: Methods and Processes*. London: Office of Health Economics; 2011.
13. Thokala P. *Multi Criteria Decision Analysis for Health Technology Assessment: Report by the Decision Support Unit*. Sheffield, England: University of Sheffield; 2011.
14. Kanavos P, Angelis A. *Multiple Criteria Decision Analysis for Value Based Assessment of New Medical Technologies: A Conceptual Framework*. London: London School of Economics and Political Science; 2013.
15. Sussex J, Rollet P, Garau M, Schmitt C, Kent A, Hutchings A. *Multi-Criteria Decision Analysis to Value Orphan Medicines*. London: Office of Health Economics; 2013.
16. 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.
17. 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(4):345–65.
18. Angelis A, Kanavos P, Montibeller G. Resource allocation and priority setting in health care: a multi-criteria decision analysis problem of value? *Global Policy*. 2017;8(Suppl. 2): 76–83.
19. Tervonen T, Naci H, van Valkenhoef G, et al. Applying multiple criteria decision analysis to comparative benefit-risk assessment: choosing among statins in primary prevention. *Med Decis Making*. 2015;35(7):859–71.
20. Ijzerman MJ, van Til JA, Snoek GJ. Comparison of two multi-criteria decision techniques for eliciting treatment preferences in people with neurological disorders. *Patient*. 2008;1(4):265–72.
21. Danner M, Hummel JM, Volz F, et al. Integrating patients' views into health technology assessment: analytic hierarchy process (AHP) as a method to elicit patient preferences. *Int J Technol Assess Health Care*. 2011;27(4):369–75.
22. Angelis A, Kanavos P. Multiple criteria decision analysis (MCDA) for evaluating new medicines in health technology assessment and beyond: the Advance Value Framework. *Soc Sci Med*. 2017;188:137–56.
23. Angelis A, Kanavos P. 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. *Pharmacoeconomics*. 2016;34(5):435–46.
24. Angelis A, Montibeller G, Hochhauser D, Kanavos P. Multiple criteria decision analysis in the context of health technology assessment: a simulation exercise on metastatic colorectal cancer with multiple stakeholders in the English setting. *BMC Med Inform Decis Mak*. 2017;17(1):149.
25. Phillips L. Decision conferencing. In: Edwards W, Miles RF Jr, von Winterfeldt D, eds. *Advances in Decision Analysis: From Foundations to Applications*. Cambridge, England: Cambridge University Press; 2007. p 375–99.
26. Phillips LD, Phillips MC. Facilitated work groups: theory and practice. *J Oper Res Soc*. 1993;44(6):533–49.
27. Franco LA, Montibeller G. Facilitated modelling in operational research. *Eur J Oper Res*. 2010;205(3):489–500.
28. von Winterfeldt D, Edwards W. *Decision Analysis and Behavioral Research*. Cambridge, England: Cambridge University Press; 1986.
29. Keeney RL, Raiffa H. *Decisions With Multiple Objectives: Preferences and Value Trade-Offs*. Cambridge, England: Cambridge University Press; 1993.
30. Bana e Costa CA, Ensslin L, Corrêa ÉC, Vansnick JC. Decision support systems in action: integrated application in a multicriteria decision aid process. *Eur J Oper Res*. 1999;113(2):315–35.
31. Tannock IF, de Wit R, Berry WR, et al; TAX 327 Investigators. Docetaxel plus prednisone or mitoxantrone plus prednisone for advanced prostate cancer. *N Engl J Med*. 2004;351(15):1502–12.
32. Berthold DR, Pond GR, Soban F, de Wit R, Eisenberger M, Tannock IF. Docetaxel plus prednisone or mitoxantrone plus prednisone for advanced prostate cancer: updated survival in the TAX 327 study. *J Clin Oncol*. 2008;26(2):242–5.
33. de Bono JS, Oudard S, Ozguroglu M, et al; TROPIC Investigators. Prednisone plus cabazitaxel or mitoxantrone for metastatic castration-resistant prostate cancer progressing after docetaxel treatment: a randomised open-label trial. *Lancet*. 2010;376(9747):1147–54.
34. de Bono JS, Logothetis CJ, Molina A, et al; COU-AA-301 Investigators. Abiraterone and increased survival in metastatic prostate cancer. *N Engl J Med*. 2011;364(21): 1995–2005.
35. Scher HI, Fizazi K, Saad F, et al; AFFIRM Investigators. Increased survival with enzalutamide in prostate cancer after chemotherapy. *N Engl J Med*. 2012;367(13):1187–97.
36. Bianchini D, Lorente D, Rodriguez-Vida A, et al. Antitumour activity of enzalutamide (MDV3100) in patients with metastatic castration-resistant prostate cancer (CRPC)

- pre-treated with docetaxel and abiraterone. *Eur J Cancer*. 2014;50(1):78–84.
37. Loriot Y, Bianchini D, Ileana E, et al. Antitumour activity of abiraterone acetate against metastatic castration-resistant prostate cancer progressing after docetaxel and enzalutamide (MDV3100). *Ann Oncol*. 2013;24(7):1807–12.
 38. National Institute for Health and Care Excellence. *Abiraterone for Castration-Resistant Metastatic Prostate Cancer Previously Treated With a Docetaxel-Containing Regimen* (Technology Appraisal Guidance 259). Manchester, England: National Institute for Health and Care Excellence; 2012.
 39. National Institute for Health and Care Excellence. *Enzalutamide for Metastatic Hormone-Relapsed Prostate Cancer Previously Treated with a Docetaxel-Containing Regimen* (Technology Appraisal Guidance 316). Manchester, England: National Institute for Health and Care Excellence; 2014.
 40. National Institute for Health and Care Excellence. *Cabazitaxel for Hormone-Refractory Metastatic Prostate Cancer Previously Treated With a Docetaxel-Containing Regimen* (Technology Appraisal Guidance 255). Manchester, England: National Institute for Health and Care Excellence; 2012.
 41. TLV. Designation 2775/2013. Tandvårds- och läkemedelsförmånsverket; 2014.
 42. TLV. Designation 4774/2014. Tandvårds- och läkemedelsförmånsverket; 2015.
 43. Keeney RL. *Value-Focused Thinking: A Path to Creative Decision Making*. Cambridge, England: Harvard University Press; 1992.
 44. Franco LA, Montibeller G. Problem structuring for multi-criteria decision analysis interventions. Available from: <https://onlinelibrary.wiley.com/doi/full/10.1002/9780470400531.eorms0683>
 45. Bana e Costa C, Beinat E. *Model-Structuring in Public Decision-Aiding*. London: London School of Economics and Political Science; 2005.
 46. Fizazi K, Scher HI, Molina A, et al; COU-AA-301 Investigators. Abiraterone acetate for treatment of metastatic castration-resistant prostate cancer: final overall survival analysis of the COU-AA-301 randomised, double-blind, placebo-controlled phase 3 study. *Lancet Oncol*. 2012; 13(10):983–92.
 47. Connock M, Cummins E, Shyangdan D, Hall B, Grove A, Clarke A. *Abiraterone Acetate for the Treatment of Metastatic, Castrate-Resistant Prostate Cancer Following Previous Cytotoxic Chemotherapy: A Single Technology Appraisal*. Coventry, England: Warwick Evidence, University of Warwick; 2011.
 48. Riemsma R, Ramaekers B, Tomini F, et al. *Abiraterone for the Treatment of Chemotherapy Naïve Metastatic Castration-Resistant Prostate Cancer: A Single Technology Appraisal*. North Yorkshire, England: Kleijnen Systematic Reviews Ltd; 2013.
 49. Collins R, Fenwick E, Trowman R, et al. A systematic review and economic model of the clinical effectiveness and cost-effectiveness of docetaxel in combination with prednisone or prednisolone for the treatment of hormone-refractory metastatic prostate cancer. *Health Technol Assess*. 2007;11(2):1–179.
 50. Kearns B, Jones LM, Stevenson M, Littlewood C. Cabazitaxel for the second-line treatment of metastatic hormone-refractory prostate cancer: a NICE single technology appraisal. *Pharmacoeconomics*. 2013;31(6):479–88.
 51. Burström K, Johannesson M, Diderichsen F. Swedish population health-related quality of life results using the EQ-5D. *Qual Life Res*. 2001;10(7):621–35.
 52. Jevtana [package insert]. Canary Wharf, London: European Medicines Agency; 2016.
 53. Xtandi [package insert]. Canary Wharf, London: European Medicines Agency; 2016.
 54. Zytiga [package insert]. Canary Wharf, London: European Medicines Agency; 2016.
 55. World Health Organization Collaborating Centre for Drug Statistics Methodology. ATC/DDD Index 2018. Available from: http://www.whocc.no/atc_ddd_index/
 56. US National Institutes of Health. Available from: ClinicalTrials.gov
 57. BNF Publications. BNF 69. The authority on the selection and use of medicines. Available from: <https://rudiapt.files.wordpress.com/2017/11/british-national-formulary-69.pdf>
 58. Bana e Costa C, Vansnick J. The MACBETH approach: basic ideas, software, and an application. In: Meskens N, Roubens M, eds. *Advances in Decision Analysis*. Dordrecht, Netherlands: Springer; 1999. p 131–57.
 59. Keeney RL. Decision analysis: an overview. *Oper Res*. 1982;30(5):803–38.
 60. Belton V, Stewart TJ. *Multiple Criteria Decision Analysis: An Integrated Approach*. Dordrecht, Netherlands: Kluwer Academic; 2002.
 61. Phillips LD. A theory of requisite decision models. *Acta Psychol*. 1984;56:29–48.
 62. Phillips LD, Bana e Costa CA. Transparent prioritisation, budgeting and resource allocation with multi-criteria decision analysis and decision conferencing. *Ann Oper Res*. 2007;154(1):51–68.
 63. Bana e Costa CA, Vansnick JC. MACBETH—an interactive path towards the construction of cardinal value functions. *Int Trans Oper Res*. 1994;1(4):489–500.
 64. Bana e Costa CA, De Corte JM, Vansnick JC. MACBETH. *Int J Inform Technol Decis Mak*. 2012;11(2): 359–87.
 65. Bana e Costa CA, Vansnick JC. Applications of the MACBETH approach in the framework of an additive aggregation model. *J Multicriteria Decis Anal*. 1997;6(2):107–14.
 66. Bana e Costa CA, Corrêa ÉC, De Corte JM, Vansnick JC. Facilitating bid evaluation in Public call for tenders: a socio-technical approach. *Omega*. 2002;30(3):227–42.
 67. Bana e Costa CA, Oliveira MD. A multicriteria decision analysis model for faculty evaluation. *Omega*. 2012;40(4): 424–36.
 68. Bana e Costa C, Lourenço J, Oliveira M, Bana e Costa J. A socio-technical approach for group decision support in

- public strategic planning: the Pernambuco PPA Case. *Group Decis Negot.* 2014;23(1):5–29.
69. Bana e Costa C, De Corte J, Vansnick J. On the mathematical foundations of MACBETH. In: Greco S, Ehrgott M, Figueira J, eds. *Multiple Criteria Decision Analysis: State of the Art Surveys*. New York: Springer; 2016. p 421–63.
 70. Keeney RL. Common mistakes in making value trade-offs. *Oper Res.* 2002;50(6):935–45.
 71. Bana e Costa C, De Corte J, Vansnick J.M-MACBETH website. Available from: http://www.m-macbeth.com/help/pdf/M-MACBETH%203.0.0%20Users%20Guide_BETA.pdf
 72. Fasolo B, Bana e Costa CA. Tailoring value elicitation to decision makers' numeracy and fluency: expressing value judgments in numbers or words. *Omega.* 2014;44:83–90.
 73. Broekhuizen H, Groothuis-Oudshoorn CG, 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(5):445–55.
 74. Tervonen T, van Valkenhoef G, Buskens E, Hillege HL, Postmus D. A stochastic multicriteria model for evidence-based decision making in drug benefit-risk analysis. *Stat Med.* 2011;30(12):1419–28.
 75. Marsh K, Ijzerman M, Thokala P, et al. Multiple criteria decision analysis for health care decision making-emerging good practices: report 2 of the ISPOR MCDA Emerging Good Practices Task Force. *Value Health.* 2016;19(2):125–37.
 76. Peacock SJ, Richardson JRJ, 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(4):897–910.
 77. Jansen JP, Fleurence R, Devine B, et al. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 1. *Value Health.* 2011;14(4):417–28.
 78. Naci H, Brugts JJ, Fleurence R, Tsoi B, Toor H, Ades AE. Comparative benefits of statins in the primary and secondary prevention of major coronary events and all-cause mortality: a network meta-analysis of placebo-controlled and active-comparator trials. *Eur J Prev Cardiol.* 2013;20(4):641–57.
 79. Naci H, Brugts JJ, Fleurence R, Ades AE. Comparative effects of statins on major cerebrovascular events: a multiple-treatments meta-analysis of placebo-controlled and active-comparator trials. *QJM.* 2013;106(4):299–306.
 80. Naci H, Brugts J, Ades T. Comparative tolerability and harms of individual statins: a study-level network meta-analysis of 246 955 participants from 135 randomized, controlled trials. *Circ Cardiovasc Qual Outcomes.* 2013;6(4):390–9.
 81. van Valkenhoef G, Tervonen T, Zhao J, de Brock B, Hillege HL, Postmus D. Multicriteria benefit-risk assessment using network meta-analysis. *J Clin Epidemiol.* 2012;65(4):394–403.
 82. Pontén J, Rönholm G, Skiold P. *PPRI Pharma Profile Sweden 2017*. Stockholm, Sweden: Pharmaceutical Pricing and Reimbursement Information; 2017.
 83. Morton A. Treacle and smallpox: two tests for multicriteria decision analysis models in health technology assessment. *Value Health.* 2017;20(3):512–5.
 84. Marsh KD, Sculpher M, Caro JJ, Tervonen T. The use of MCDA in HTA: great potential but more effort is needed. *Value Health.* 2018;21(4):394–7.
 85. Kirkwood CW. *Strategic Decision Making: Multiobjective Decision Analysis With Spreadsheets*. Belmont (CA): Duxbury Press; 1997.
 86. Bana e Costa CA, Carnero MC, Oliveira MD. A multi-criteria model for auditing a predictive maintenance programme. *Eur J Oper Res.* 2012;217(2):381–93.
 87. Dyer J. MAUT—multiattribute utility theory. In: Figueira J, Greco S, Ehrgott M, eds. *Multiple Criteria Decision Analysis: State of the Art Surveys*. New York: Springer; 2005. p 265–96.
 88. Nord E, Daniels N, Kamlet M. QALYs: some challenges. *Value Health.* 2009;12(Suppl. 1):S10–S5.
 89. Dyer J, Sarin R. Measurable multiattribute value functions. *Oper Res.* 1979;27(4):810–22.
 90. Chongtrakul P, Sumpradit N, Yoongthong W. ISaE and the evidence-based approach for essential medicines selection in Thailand. *Essent Drugs Monit.* 2005;34:18–9.
 91. Rodrigues TC, Montibeller G, Oliveira MD, Bana E Costa CA. Modelling multicriteria value interactions with reasoning maps. *Eur J Oper Res.* 2017;258(3):1054–71.
 92. Claxton K. Should multi-criteria decision analysis (MCDA) replace cost effectiveness analysis (CEA) for evaluation of health care coverage decisions? Paper presented at: ISPOR 16th Annual European Congress; November 2–4, 2013; Dublin, Ireland.
 93. Schaffer KS, Sussex J, Devlin N, Walker A. Local health care expenditure plans and their opportunity costs. *Health Policy.* 2015;119(9):1237–44.
 94. Claxton K, Martin S, Soares M, et al. Methods for the estimation of the National Institute for Health and care excellence cost-effectiveness threshold. *Health Technol Assess.* 2015;19(14):1–503.