

## PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/212262>

Please be advised that this information was generated on 2021-09-18 and may be subject to change.



ScienceDirect

Contents lists available at [sciencedirect.com](http://sciencedirect.com)  
Journal homepage: [www.elsevier.com/locate/jval](http://www.elsevier.com/locate/jval)

Methodology

## Multicriteria Decision Analysis to Support Health Technology Assessment Agencies: Benefits, Limitations, and the Way Forward



Rob Baltussen, PhD,<sup>1,\*</sup> Kevin Marsh, PhD,<sup>2</sup> Praveen Thokala, PhD,<sup>3</sup> Vakaramoko Diaby, PhD,<sup>4</sup> Hector Castro, PhD,<sup>5</sup> Irina Cleemput, PhD,<sup>6</sup> Martina Garau, PhD,<sup>7</sup> Georgi Iskrov, PhD,<sup>8,9</sup> Alireza Olyaeemanesh, PhD,<sup>10</sup> Andrew Mirelman, PhD,<sup>11</sup> Mohammedreza Mobinizadeh, PhD,<sup>10</sup> Alec Morton, PhD,<sup>12</sup> Michele Tringali, PhD,<sup>13</sup> Janine van Til, PhD,<sup>14</sup> Joice Valentim, PhD,<sup>15</sup> Monika Wagner, PhD,<sup>16</sup> Sitaporn Youngkong, PhD,<sup>17</sup> Vladimir Zah, PhD,<sup>18</sup> Agnes Toll, MSc,<sup>1</sup> Maarten Jansen, MSc,<sup>1</sup> Leon Bijlmakers, PhD,<sup>1</sup> Wija Oortwijn, PhD,<sup>1</sup> Henk Broekhuizen, PhD<sup>1</sup>

<sup>1</sup>Radboud University Medical Center, Nijmegen, The Netherlands; <sup>2</sup>Evidera, London, England, UK; <sup>3</sup>University of Sheffield, Sheffield, England, UK; <sup>4</sup>Florida Agricultural and Mechanical University, Tallahassee, FL, USA; <sup>5</sup>Management Sciences for Health, Arlington, VA, USA; <sup>6</sup>Belgian Health Care Knowledge Centre, Brussels, Belgium; <sup>7</sup>Office of Health Economics, London, England, UK; <sup>8</sup>Medical University of Plovdiv, Plovdiv, Bulgaria; <sup>9</sup>Institute for Rare Diseases, Plovdiv, Bulgaria; <sup>10</sup>Tehran University of Medical Sciences, Tehran, Iran; <sup>11</sup>University of York, York, England, UK; <sup>12</sup>University of Strathclyde, Glasgow, Scotland; <sup>13</sup>Lombardia Regional Health Directorate, Milan, Italy; <sup>14</sup>University of Twente, Enschede, The Netherlands; <sup>15</sup>Roche, Basel, Switzerland; <sup>16</sup>LASER Analytica, Montreal, Canada; <sup>17</sup>Mahidol University, Bangkok, Thailand; <sup>18</sup>ZRx Outcomes Research Inc, Mississauga, Canada.

### ABSTRACT

**Objective:** Recent years have witnessed an increased interest in the use of multicriteria decision analysis (MCDA) to support health technology assessment (HTA) agencies for setting healthcare priorities. However, its implementation to date has been criticized for being “entirely mechanistic,” ignoring opportunity costs, and not following best practice guidelines. This article provides guidance on the use of MCDA in this context.

**Methods:** The present study was based on a systematic review and consensus development. We developed a typology of MCDA studies and good implementation practice. We reviewed 36 studies over the period 1990 to 2018 on their compliance with good practice and developed recommendations. We reached consensus among authors over the course of several review rounds.

**Results:** We identified 3 MCDA study types: qualitative MCDA, quantitative MCDA, and MCDA with decision rules. The types perform differently in terms of quality, consistency, and transparency of recommendations on healthcare priorities. We advise HTA agencies to always include a deliberative component. Agencies should, at a minimum, undertake qualitative MCDA. The use of quantitative MCDA has additional benefits but also poses design challenges. MCDA with decision rules, used by HTA agencies in The Netherlands and the United Kingdom and typically referred to as *structured deliberation*, has the potential to further improve the formulation of recommendations but has not yet been subjected to broad experimentation and evaluation.

**Conclusion:** MCDA holds large potential to support HTA agencies in setting healthcare priorities, but its implementation needs to be improved.

**Keywords:** HTA agencies, multicriteria decision analysis, priority setting, value framework.

VALUE HEALTH. 2019; 22(11):1283–1288

## Introduction

Recent years have witnessed an increased interest in the use of multicriteria decision analysis (MCDA) to support health technology assessment (HTA) agencies in setting healthcare priorities.<sup>1–6</sup> MCDA offers decision makers a structured way to include the different values that society holds. The term *values* here refers to both the therapeutic benefits of a technology for patients and their broader social impact.<sup>7,8</sup> For example, decision makers may value technologies that not only maximize population health but also reduce health inequalities or protect people against the

impoverishing effects of ill health.<sup>9</sup> However, the way that MCDA has been implemented to date has been criticized for being “entirely mechanistic,”<sup>10</sup> ignoring opportunity costs,<sup>11,12</sup> and paying insufficient attention to best practice guidelines.<sup>11–14</sup> Subsequently, some HTA agencies and scholars have rejected its use.<sup>10,11,15,16</sup>

This article provides guidance on the use of MCDA by HTA agencies. We present a typology of MCDA studies, a review of studies over the period 1990 to 2018 for illustrative purposes, and a critical assessment of the various study types (second through fifth sections). We judge the ability of the study types to improve

\* Address correspondence to: Rob Baltussen, PhD, PO Box 9101, 6500 HB Nijmegen, The Netherlands. Email: [rob.baltussen@radboudumc.nl](mailto:rob.baltussen@radboudumc.nl)

the recommendations of HTA agencies, in terms of their quality (by taking into account all relevant [stakeholder] values and making appropriate trade-offs between them and by capturing opportunity costs), their consistency (by repeatedly considering the same values), and their transparency (by being explicit on the selection of values and the performance of technologies on these values).<sup>4,5,7</sup> Together, this ultimately improves the legitimacy of recommendations.<sup>7</sup> Finally, we provide recommendations on the use of MCDA in HTA decision making (see the “Discussion” section).

This article addresses governmental HTA agencies but is also relevant for countries that have not (yet) established such an agency. We assume a model in which an agency has installed an (appraisal) committee that formulates recommendations for the Ministry of Health on technology coverage decisions. We consider this task of HTA agencies as an intrinsically complex and value-laden political process.<sup>17–21</sup> Society, including relevant stakeholders such as healthcare providers, patients, citizens, funders, and decision makers, have different interests and may reasonably disagree on what values should be used to guide priority setting.<sup>20</sup> Because national governments are accountable to the populations they serve, they should be concerned with establishing legitimate decision-making processes that take these values into account.<sup>22</sup> Stakeholder deliberation is considered an essential component to achieve such legitimacy.<sup>7,20</sup>

The paper follows up on the recent ISPOR MCDA Emerging Good Practices Task Force on the use of MCDA in healthcare<sup>4,5</sup> by being specific on its application to HTA. It is written by 23 MCDA and HTA experts, and is the result of intensive discussions using various review rounds. It can be considered as a consensus statement on the benefits and limitations of the use of MCDA for HTA agencies, and its way forward.

## A Typology of MCDA Studies

MCDA is defined as “an umbrella term to describe a collection of formal approaches which seek to take explicit account of multiple criteria in helping individuals or groups exploring decisions that matter.”<sup>4,23</sup> Any MCDA involves at least 3 steps: defining the decision problem, selecting the criteria that reflect relevant values, and constructing the performance matrix.<sup>4</sup> The performance matrix is a central element and, when applied to HTA, typically includes a set of generic criteria that are relevant to many technologies. The performance matrix presents an assessment of each technology against each of these criteria using

descriptive information, such as natural units (eg, number of deaths), categories (eg, targeted age group), summary measures of health (eg, quality-adjusted life-years [QALYs]), or descriptive text (eg, perceived role of own responsibility). (An example can be found in Fig. 1).<sup>24</sup> The committee evaluates the performance matrix before formulating a recommendation.<sup>24</sup> They may rely on the criteria included in the performance matrix and, if applicable, include other considerations specific to the technology under scrutiny.<sup>7</sup>

We distinguish “qualitative MCDA,” “quantitative MCDA,” and “MCDA with decision rules,” depending on the way the performance matrix is used. By reviewing the literature on the use of MCDA in HTA between 1990 and 2018, we identified 36 studies and classified these accordingly (review details are presented in Appendix A in the Supplemental Materials found at <https://doi.org/10.1016/j.jval.2019.06.014>). Our review identified 1 qualitative MCDA study, 35 quantitative MCDA studies, and no studies using MCDA with decision rules. We nevertheless included the latter as a distinct study type because of its use by some HTA agencies.

## Qualitative MCDA

In qualitative MCDA, the committee makes a judgment on the overall value of a technology by deliberating on its performance regarding explicitly defined criteria (ie, it makes a qualitative interpretation of the performance matrix; Fig. 1).<sup>7</sup> This approach is also referred to as *partial MCDA*<sup>4,5</sup> or the *balance sheet method*.<sup>25</sup> The qualitative MCDA study included in our review pertains to the development of the health benefit package in Thailand.<sup>26</sup>

The distinctive feature of qualitative MCDA that makes it different from intuitive prioritization (without any specific method) is that it uses explicit criteria, including the technologies’ performance on these criteria. This has several implications for the quality, consistency, and transparency of recommendations.

First, the use of explicit criteria improves the quality of recommendations as it fosters in-depth consideration of the criteria, including the available evidence, and it provides structure to deliberative discussions of the committee.<sup>27</sup> This arguably reduces the committee’s cognitive load of simultaneously processing information on otherwise implicit criteria. However, the cognitive load may still be extensive, especially when it involves the simultaneous evaluation of multiple technologies requiring complex trade-offs between criteria. In addition, qualitative MCDA carries the risk that certain stakeholders dominate the deliberations,

**Figure 1.** Interpretation of performance matrix in qualitative multicriteria decision analysis. Severity of disease is shown as a 4-star scale, with more stars indicating a more severe disease.

Technologies	Criteria			
	Effectiveness (quality adjusted life years)	Severity of disease	Disease of the poor	Age
Antiretroviral treatment in HIV/AIDS	100	●●●●	√	15 years and older
Treatment of childhood pneumonia	200	●●●●	√	0-14 years
Inpatient care for acute schizophrenia	10	●●		15 years and older
Plastering for simple fractures	200	●		all

especially in contexts with unbalanced power relationships.<sup>24,25,27,28</sup> This may reduce the quality of recommendations unless mechanisms to minimize dominance are installed.<sup>29</sup> Furthermore, it depends on the included criteria whether qualitative MCDA facilitates a comparison with alternative uses of resources and thereby captures opportunity costs.

Second, if the same set of explicit criteria is repeatedly used in other evaluations, qualitative MCDA improves the consistency of eventual recommendations. Yet this consistency can be limited because a committee may judge the importance of criteria across evaluations differently. This can be overcome by making the argumentation underlying a recommendation explicit: as such, these argumentations can be referred to, and applied, in the formulation of recommendations on other technologies.

Third, the use of explicit criteria improves the transparency of recommendations. Full transparency would also require that the argumentation for making a recommendation is made public. However, this may not always be feasible (eg, in countries with limited tradition of transparency and accountability in public decision making).

## Quantitative MCDA

Quantitative MCDA (also labeled *full MCDA*<sup>4,5</sup>) uses a value measurement model to interpret the performance matrix, followed by deliberation. This approach includes 5 further steps, in addition to the 3 steps described earlier to construct the performance matrix.<sup>4</sup> First, stakeholders' preferences are elicited to specify a value function for each criterion, which translates a technology's performance on that criterion into a score (eg, between 0 and 100). Second, stakeholders' preferences regarding the relative importance of criteria are measured using criterion weights. Various preference elicitation techniques such as analytic hierarchy process or discrete-choice experiments are available for this.<sup>4,5</sup> Group preferences are often modeled by taking the mean criterion weights and scores across respondents. Third, a so-called value measurement model is used, which typically multiplies scores by the relative weight of that criterion, to sum the weighted scores and obtain an overall value for each technology. Technologies are ranked on the basis of these overall values (an example is provided in Fig. 2).<sup>4,5</sup> Fourth, uncertainty analysis is performed to understand the level of robustness of the results. Fifth, the committee deliberates on this rank ordering of technologies, allowing a flexible interpretation of the results; that

is, its members can put forward and discuss (aspects of) criteria that were not (fully) captured in the performance matrix (eg, on complex consideration such as "own responsibility"). This step may lead to changes in the ordering of technologies. With 35 studies, quantitative MCDA is the most common study type in our review.

Quantitative MCDA has several benefits in comparison to qualitative MCDA. First, the use of a value measurement model reduces the cognitive load of processing several criteria simultaneously and the risk of dominant participants influencing the deliberations. These aspects further contribute to the quality of recommendations. Second, the use of criteria scores and weights further improves the consistency of recommendations, if these scores and weights are also used for the evaluation of other technologies. Third, when these aspects are also communicated to the public, it further enhances the transparency of recommendations.

These benefits of using explicit weights are especially relevant for HTA agencies in certain contexts. If an agency operates in a country with limited tradition of transparency and accountability in public decision making, this may raise trust in its decision making. Also, the use of quantitative MCDA can be instrumental if an agency operates in a country with a backlog of technologies waiting for appraisal and insufficient HTA capacity for more detailed evaluations. We here refer to the case of the Colombian HTA agency in 2012 to 2013, which faced the task of assessing hundreds of technologies with very limited capacity.<sup>30</sup>

There are also various limitations to quantitative MCDA, which typically relate to its implementation rather than fundamental problems (for an overview of limitations by study, see [Appendix Table A1](#) in the Supplemental Materials) and which may compromise the quality of recommendations. Here we focus on 5 of them. First, although quantitative MCDA should always include a deliberative component allowing a committee to make a flexible interpretation of results, only 10 out of 35 studies reported such a deliberation. This has likely led to the neglect of additional considerations that are specific to the technology under scrutiny and, subsequently, confounded recommendations. This suboptimal practice seems to have led HTA agencies in the United Kingdom and The Netherlands to explicitly reject quantitative MCDA. An expert meeting conducted by the National Institute for Health and Care Excellence (NICE) in the United Kingdom in 2012 concluded,

The majority of participants agreed that once the committee has decided what the plausible incremental cost-effectiveness ratio is, the decision-making process should remain deliberative and flexible, rather

**Figure 2.** Interpretation of performance matrix in quantitative multicriteria decision analysis. Preference scores for effectiveness are related to its values, following a linear scale. For disease of the poor, if the technology targets a disease of the poor, it scores 100, otherwise 0. Preference scores for severity of disease are scaled between 0 and 100 in proportion to their bullets in the table. Assuming decision-makers have a preference to treat young people over old, 0 to 14 years receives a score of 100, 15 years and older a score of 0, and all ages a score of 50. Preference scores are presented here for illustrative purposes only and are arbitrary.

Technologies	Effectiveness	Severity of disease	Disease of the poor	Age	Overall value
Antiretroviral treatment in HIV/AIDS	50	100	100	0	70
Treatment of childhood Pneumonia	100	100	100	100	100
Inpatient care for acute Schizophrenia	5	50	0	0	7
Plastering for simple fractures	100	25	0	50	48
Weights	40	10	40	10	

than moving towards a fully quantitative (or algorithmic) approach. . . . The vast majority of participants did not recommend that NICE should attempt to assign weights to the additional criteria, suggesting that flexible deliberation is important rather than stringent rule.(...) <sup>16,31</sup>

An expert meeting in The Netherlands, organized by the National Health Care Institute (ZIN) in 2015, drew a similar conclusion, arguing that deliberation should always be part of its process to formulate recommendations. <sup>15</sup>

Of the 10 studies that did include deliberation, this component changed the initial rank order of technologies in only 3 studies. This suggests that end users agree with the results from the value measurement model or that they rely solely on its results. <sup>28</sup> If the latter is true, this indicates the need to organize adequate deliberative components in quantitative MCDA.

Second, 25 out of 35 studies used an additive value measurement model, which embeds the preferential independence assumption (ie, how people appreciate performance on one criterion does not depend on the performance on other criteria). <sup>12,13,28</sup> This assumption does not always hold. For example, a technology that does not improve population health has no value, irrespective of whether it targets a severe or rare disease. This is illustrated by a quantitative MCDA on anticancer drugs in South Korea, in which stakeholders assigned a weight of 22% to the criterion “clinical benefit.” <sup>32</sup> It demonstrates that a technology that is ineffective but performs well on all other criteria can still obtain a high value. Whereas quantitative MCDA would label this technology as high priority, it is clearly inappropriate. Value measurement models could include interaction weights between criteria, but measuring these requires complex elicitation approaches that are cognitively demanding.

Third, 29 of 36 studies included costs as a criterion in their value measurement model and subsequently applied the “maximizing value” allocation rule. Here, the aggregate value of a technology (ie, derived from the value functions of the individual criteria) includes the value of the related costs (derived from the value function of the cost criterion, in which higher costs are related to lower values). Technologies with the highest overall value are then considered priorities. The approach requires respondents to derive value functions for all criteria including the cost criterion and provide weights for the value function of cost in relation to that of the other criteria. However, in practice, it is unrealistic to assume that individuals can adequately fulfill this task. It is unlikely that they are aware of health budget constraints and alternative ways of using resources, and their responses therefore do not adequately capture the opportunity costs of alternatives. <sup>5,12,33</sup> This may result in a confounded ranking of technologies (see [Appendix B](#) in the Supplemental Materials for a numerical example).

For this reason, several authors have instead proposed the use of the traditional “cost-per-value” allocation rule. <sup>5,12,34</sup> Here, the costs of a technology are divided by the aggregate value of other criteria. Subsequently, technologies are rank ordered on the basis of their cost-per-value ratio, and technologies with the lowest ratios are considered priorities. To achieve an optimal allocation of resources and adequately take into account opportunity costs, HTA agencies are then advised to fund technologies according to this ranking until the budget is exhausted. Or, as an alternative approach, the cost-per-value of a technology can be compared with a threshold. However, in reality, budget constraints are seldom explicit, and cost-per-value thresholds are typically unknown. <sup>35</sup> This means that quantitative MCDA can provide only a ranking of technologies and cannot be explicit on whether technologies are providing value for money. This approach may, in the absence of

recognition of opportunity costs, result in a suboptimal allocation of resources.

Closely related, 20 studies included “cost-effectiveness,” such as cost per QALY, as a criterion in the value measurement model. The approach requires respondents to derive a value function and weights for the cost-effectiveness criterion in relation to that of the other criteria, thereby reflecting the opportunity costs of alternatives. As reasoned earlier, it is unrealistic to assume that individuals can adequately perform this task.

Fourth, 25 out of 35 studies involved double counting of 1 or more criteria. This indicates problems in the structuring phase of the MCDA value measurement model. <sup>36</sup>

Fifth, 2 out of 35 studies did not use preference-based techniques such as an analytic hierarchy process or discrete-choice experiments for eliciting scores and weights but applied simple direct rating methods such as point allocation. These studies risk eliciting scores and weights that are subject to framing bias, as criteria and their performance ranges are not explicitly traded off. <sup>12</sup> In addition, these studies often provide no or only qualitative descriptions of performance ranges, and respondents may interpret these ranges differently. <sup>12</sup>

## MCDA With Decision Rules

In MCDA with decision rules, the committee interprets the performance matrix with a set of simple rules. These rules guide them in making trade-offs between criteria, which can be quantitative or qualitative in nature. Some HTA agencies follow this approach, defining the relationship between cost-effectiveness and other criteria. For example, ZIN in The Netherlands appraises the cost-effectiveness of technologies in relation to the severity of the condition. Technologies that target mild conditions (ie, below 0.4 on a burden of disease scale from 0 to 1) should cost less than €20 000 per QALY to receive an initial positive recommendation for reimbursement. Technologies targeting severe and very severe conditions (ie, between 0.4 and 0.7 and greater than 0.7) may cost up to €50 000 and €80 000 per QALY, respectively. Subsequently, ZIN evaluates in a deliberative process whether other criteria affect the initial recommendation and reaches a final recommendation. <sup>37</sup> In the United Kingdom, NICE has issued decision rules on the relationship between cost-effectiveness and other criteria:

Above a most plausible ICER [incremental cost-effectiveness ratio] of £20,000 per QALY gained, judgements about the acceptability of the technology as an effective use of NHS resources will specifically take account of the following factors: The degree of certainty around the ICER . . . , the innovative nature of the technology . . . , the technology meets the criteria for special consideration as a ‘life-extending treatment at the end of life’ . . . , and aspects that relate to non-health objectives of the NHS. <sup>38</sup>

In its highly specialized technology program for very rare diseases, NICE raised the threshold to £100 000 to £300 000 per QALY gained. This increased threshold reflects the fact that NICE assigns a quantitative weight to the treatment of these diseases. <sup>39</sup>

We label the NICE and ZIN approaches as MCDA, whereas we acknowledge that they are usually referred to as structured deliberation. <sup>37,38</sup> We do so as the approaches fit within the MCDA definition, as provided earlier, to “take explicit account of multiple criteria in helping individuals or groups exploring decisions that matter.” We hereby wish to bridge the artificial gap between so-called deliberative and MCDA approaches and to stimulate the debate on how multiple criteria can best be taken into account.

How does MCDA with decision rules compare with quantitative MCDA? First, the approach can incorporate the principle of opportunity costs, if cost-effectiveness is used as a central criterion and its threshold is known. It therefore improves on quantitative MCDA, which does not capture opportunity costs and could therefore be considered less suitable for priority setting. Second, the approach as applied by ZIN and NICE includes only a limited number of criteria in its decision rules. This contrasts with the multiple (often 10 or more) criteria that are typically included in quantitative MCDA. Although this is not a necessary difference between the MCDA designs, experience so far indicates that this allows MCDA with decision rules to more rigorously define and assess the most important criteria. A disadvantage is that MCDA with decision rules may involve more deliberation around the remaining criteria that are not included in decision rules. Deliberation can take more time than value measurement and may also lead to less consistent and transparent recommendations if not well documented.

## Discussion

The core challenge for HTA agencies is to optimize the quality, consistency, and transparency of their recommendations for priority setting. This article shows that various MCDA types perform differently with regard to these aspects. Here, we provide recommendations for HTA agencies and the research community on the future use of MCDA types.

First, we advise HTA agencies to always include a deliberative component in their process of formulating recommendations. This allows the relevant committee a flexible interpretation of decision-making criteria to take into account all possible considerations that matter. Such deliberation may improve the quality of recommendations. Agencies should report these deliberations, including the considerations underlying a recommendation, to ensure the consistency and transparency of recommendations.

Second, agencies should, at a minimum, undertake qualitative MCDA. The use of explicit criteria improves the quality, consistency, and transparency of recommendations as compared with employing no specific method at all, although important challenges remain.

Third, HTA agencies may consider the use of quantitative MCDA. A number of HTA agencies have already implemented this approach<sup>40,41</sup> but base their recommendations on the value measurement model only. We recommend that they work toward the incorporation of deliberative elements into their MCDA designs in the future. Quantitative MCDA also poses other design challenges. Specifically, we advise researchers not to include “cost” or “cost-effectiveness” as criteria in the value measurement model. More generally, we advise researchers to follow good practice and indicate the potential confounding that stems from suboptimal designs.<sup>4</sup> HTA agencies should be aware of these challenges when interpreting the results. They should also be aware that quantitative MCDA does not capture opportunity costs and may thus lead to a suboptimal allocation of resources.

Fourth, HTA agencies may consider the use of MCDA with decision rules. This approach has the same potential as quantitative MCDA to improve decision making, but depending on the included number of criteria, may rely more on deliberation. It also avoids certain challenges in study design and can capture opportunity costs. The approach is now routinely used in The Netherlands and the United Kingdom (albeit named differently), and this demonstrates it is workable in practice. However, it has not been subjected to broad experimentation and evaluation, and we call for

research to demonstrate the added value of MCDA with decision rules.

Fifth, HTA agencies should ensure that the specification of MCDA (ie, in terms of value measurement model or decision rules) is legitimate and reflects societal preferences. The debate in the United Kingdom on a proposed “value-based assessment” framework demonstrates this challenge.<sup>42,43</sup> This article has not discussed how to best elicit stakeholder preferences, and we call for further debate and guidance on this topic.

This article makes a significant contribution to the literature on the use of MCDA in HTA. It follows up on the ISPOR MCDA Emerging Good Practices Task Force on the use of MCDA in healthcare<sup>4,5</sup> by providing guidance on its specific application to HTA. In addition, it provides a head-to-head comparison of the different MCDA study types, identifying the options and limitations of each approach, and providing recommendations on its use by HTA agencies. We thereby define MCDA with decision rules as a separate MCDA study type, although this approach is typically referred to as “structured deliberation,” and our review did not identify any such study. We nevertheless did so to stimulate the debate on how multiple criteria can best be taken into account.

Our recommendations should be interpreted in the context of the following aspects. First, our literature review includes only studies that are self-described MCDA. Many other studies exist that do consider multiple criteria but are not labeled as such. This may explain the small number of studies found, with most of them focusing on quantitative MCDA, only one on qualitative MCDA, and none on MCDA with decision rules. Our literature review should then also be considered as illustrative only. Second, we evaluate MCDA study types on their ability to improve the quality of recommendations. With our definition of quality, we aim to identify and differentiate the most important options and limitations of the study types. The definition is not meant to capture all aspects of quality of decision making (eg, quality of evidence or quality of stakeholder deliberation). Third, we evaluated study types in terms of transparency of forthcoming recommendations, but HTA agencies or involved stakeholders may not always aim for full transparency of their reimbursement recommendations (eg, in the case of price negotiations with providers). Fourth, in the reality of healthcare priority setting (in which decisions on technologies are typically taken for single technologies), we argue that quantitative MCDA cannot capture opportunity costs. However, in specific circumstances in which decisions are made for a complete set of technologies in the presence of a fixed budget, mathematical programming techniques can be used to develop optimal solutions.

In conclusion, MCDA holds large potential to support HTA agencies in formulating high-quality, consistent, and transparent recommendations. However, its application has often been inadequate and subject to criticism. We consider it the shared responsibility of HTA agencies, the research community, and decision makers to improve on the use of MCDA, to realize its full potential.

## Acknowledgments

This work was funded by a VICI fellowship for R.B. from the Dutch Research Council.

## Supplementary Data

Supplementary data associated with this article can be found in the online version at <https://doi.org/10.1016/j.jval.2019.06.014>.

## REFERENCES

1. 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):345–365.
2. Thokala P, Marsh K, Devlin N, et al. Multi criteria decision analysis methods in health care: current status, good practice and future recommendations. *Value Health*. 2014;17(3):A34–A34.
3. Adunlin G, Diaby V, Xiao H. Application of multicriteria decision analysis in health care: a systematic review and bibliometric analysis. *Health Expect*. 2015;18(6):1894–1905.
4. Thokala P, Devlin N, Marsh K, et al. Multiple criteria decision analysis for health care decision making—an introduction: report 1 of the ISPOR MCDA Emerging Good Practices Task Force. *Value Health*. 2016;19:1–13.
5. 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):125–137.
6. Marsh K, Thokala P, Goetghebeur M, Baltussen R, eds. *Multi-criteria Decision Analysis to Support Healthcare Decisions*. New York, NY: Springer; 2017.
7. Baltussen R, Jansen M, Bijlmakers L, et al. Value assessment frameworks for HTA agencies: the organization of evidence-informed deliberative processes. *Value Health*. 2017;20(2):256–260.
8. 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):123–152.
9. Tromp N, Baltussen R. Mapping of multiple criteria for priority setting of health interventions: an aid for decision makers. *BMC Health Serv Res*. 2012;12:454.
10. Kennedy I. *Appraising the Value of Innovation and Other Benefits: A Short Study for NICE*. London, UK: National Institute for Clinical Excellence; 2009.
11. Campillo-Artero C, Puig-Junoy J, Culyer AJ. Does MCDA trump CEA? *Appl Health Econ Health Policy*. 2018;16(2):147–151.
12. Marsh KD, Sculpher M, Caro JB, Tervonen T. The use of MCDA in HTA: great potential, but more effort needed. *Value Health Reg Issues*. 2018;21(4):394–397.
13. Morton A. Treacle and smallpox: two tests for multicriteria decision analysis models in health technology assessment. *Value Health*. 2017;20:512–515.
14. Wahlster P, Goetghebeur M, Kriza C, Niederlander C, Kolominsky-Rabas P. National Leading-Edge Cluster Medical Technologies ‘Medical Valley EMN.’ Balancing costs and benefits at different stages of medical innovation: a systematic review of multi-criteria decision analysis (MCDA). *BMC Health Serv Res*. 2015;15:262.
15. National Health Care Institute. *Expert Meeting ACP on Multi Criteria Decision Analysis*. Diemen, The Netherlands: National Health Care Institute; 2015.
16. National Institute for Health and Clinical Excellence. *Key Issues Arising From Workshop on Structured Decision Making*. Prepared by George E. London, UK: National Institute for Health and Clinical Excellence; 2012.
17. Holm S. The second phase of priority setting. Goodbye to the simple solutions: the second phase of priority setting in health care. *Br Med J*. 1998;317(7164):1000–1002.
18. Mitton C, Donaldson C. Health care priority setting: principles, practice and challenges. *Cost Eff Resour Alloc*. 2004;2(1):3.
19. Kipiriri L, Martin DK. A strategy to improve priority setting in developing countries. *Health Care Anal*. 2007;15(3):159–167.
20. Daniels N. Accountability for reasonableness. *Br Med J*. 2000;321(7272):1300–1301.
21. Abelson J, Giacomini M, Lehoux P, Gauvin FP. Bringing ‘the public’ into health technology assessment and coverage policy decisions: from principles to practice. *Health Policy*. 2007;82(1):37–50.
22. Goetghebeur M, Castro-Jaramillo H, Baltussen R, Daniels N. The art of priority setting. *Lancet*. 2017;389(10087):2368–2369.
23. Belton V, Stewart TJ. *Multi Criteria Decision Analysis: An Integrated Approach*. Boston, MA: Kluwer Academic; 2002.
24. Baltussen R, Niessen L. Priority setting of health interventions: the need for multi-criteria decision analysis. *Cost Eff Resour Alloc*. 2006;4:14.
25. Makundi E, Kipiriri L, Norheim OF. Combining evidence and values in priority setting: testing the balance sheet method in a low-income country. *BMC Health Serv Res*. 2007;7:152.
26. 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):961–970.
27. Brunetti M, Shemilt I, Pregno S, et al. GRADE guidelines: 10. Considering resource use and rating the quality of economic evidence. *J Clin Epidemiol*. 2013;66(2):140–150.
28. Devlin N, Sussex J. *Incorporating Multiple Criteria in HTA: Methods and Processes*. London, UK: Office of Health Economics; 2011.
29. Kadlec A, Friedman W. Deliberative democracy and the problem of power. *J Public Deliber*. 2007;3(1).
30. Jaramillo HE, Goetghebeur M, Moreno-Mattar O. Testing multi-criteria decision analysis for more transparent resource-allocation decision making in Colombia. *Int J Technol Assess Health Care*. 2016;32(4):307–314.
31. National Institute for Health and Clinical Excellence. *Briefing Paper for Methods Review Workshop on Structured Decision Making*. Prepared by Claxton C and Devlin N. London, UK: National Institute for Health and Clinical Excellence; 2011.
32. Kwon SH, Park SK, Byun JH, Lee EK. Eliciting societal preferences of reimbursement decision criteria for anti cancer drugs in South Korea. *Exp Rev Pharmacoecon Outcomes Res*. 2017;17(4):411–419.
33. Garau M, Devlin NJ. Using MCDA as a decision aid in health technology appraisal for coverage decisions: opportunities, challenges and unresolved questions. In: Marsh K, Goetghebeur M, Thokala P, Baltussen R, eds. *Multi-criteria Decision Analysis to Support Healthcare Decisions*. New York, NY: Springer; 2017:277–298.
34. 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.
35. Thokala P, Ochalek J, Leech AA, Tong T. Cost-effectiveness thresholds: the past, the present and the future. *Pharmacoeconomics*. 2018;36(5):509–522.
36. Marttunen M, Lienert J, Belton V. Structuring problems for multi-criteria decision analysis in practice: a literature review of method combinations. *Eur J Oper Res*. 2017;263(1):1–17.
37. *Kosteneffectiviteit in de praktijk* [Cost-effectiveness in practice]. Diemen, The Netherlands: National Health Care Institute (ZIN); 2015.
38. National Institute for Health and Care Excellence. *Guide to the Methods of Technology Appraisal*. London, UK: National Institute for Health and Care Excellence; 2013.
39. National Institute for Clinical Excellence. *Highly specialised technologies guidance*. <https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/nice-highly-specialised-technologies-guidance>. Accessed March 9, 2018.
40. Iskrov G, Miteva-Katrandzhieva T, Stefanov R. Multi-criteria decision analysis for assessment and appraisal of orphan drugs. *Front Public Health*. 2016;4:214.
41. Gulacsi L, Rotar AM, Niewada M, et al. Health technology assessment in Poland, the Czech Republic, Hungary, Romania and Bulgaria. *Eur J Health Econ*. 2014;15(suppl 1):S13–S25.
42. Kusel J. Why has value based assessment been abandoned by NICE in the UK? *Value Outcomes Spotlight*. 2015;1:22–25.
43. Chalkidou K. Evidence and values: paying for end-of-life drugs in the British NHS. *Health Econ Policy L*. 2012;7(4):393–409.