

The Potential and Pitfalls of Using Multi-criteria Decision Analysis to Support Health Technology Assessment in Europe

Janine van Til, PhD, University of Twente, Enschede, The Netherlands; Susanne Schmitz, PhD, Luxembourg Institute of Health, Strassen, Luxembourg; Michele Tringali, PhD, MD, Regione Lombardia, Milan, Italy; and Rob Baltussen, PhD, Radboudumc-Department for Health Evidence, Humlebæk, Denmark

Multi-criteria decision analysis (MCDA) has the potential to improve the structure, transparency, consistency, accountability, and validity of deliberative decision-making processes.

In most European countries, decisions to reimburse innovations in health care are made after a long and complex process of problem analysis, evidence gathering, and assessment. In most countries, evidence gathering and assessment are followed by deliberative sessions with experts from different backgrounds within and outside health care who appraise the evidence, share their perspective, and discuss the need for reimbursement of an innovation from a societal perspective. This decision process is influenced by multiple criteria. Sometimes the assessment criteria are stated explicitly by the agency; sometimes none or only a subset of criteria is made explicit. For instance, cost-effectiveness of the intervention is an important criterion in most European countries. Moreover, the severity of the disease, the safety and tolerability of the innovation, the quality of the evidence, and budget impact of reimbursement might also affect decisions in Europe.

With the growing costs of health care and the need for budget management, decision panels are increasingly asked to justify their decisions. Multi-criteria decision analysis (MCDA) is often proposed as a way to support reimbursement decision processes. MCDA is “an umbrella term to describe a collection of formal approaches, which seek to take explicit account of multiple criteria in helping individuals or groups explore decisions that matter” (Belton and Steward, 2003).¹ From a theoretical perspective, it seems very promising to use MCDA to support reimbursement decisions, as it is developed to support complex decisions driven by multiple and possibly conflicting arguments, multiple stakeholders, and in which there is no obvious “right” decision.

This article will explore the potential and pitfalls of MCDA to increase structure, transparency, consistency, and validity of deliberative decision processes, based on our experiences in Europe with introducing MCDA to health technology assessment (HTA) decision making.

The Potential of MCDA to Support HTA Although MCDA consists of a wide range of techniques, the common denominator is that all methods follow a stepwise approach to decision making.^{2,3} By following this stepwise approach, the promise of MCDA lies in improving the structure, transparency, consistency, validity, and accountability of the decisions that are made. The structure of decisions refers to the extent to which the organizational body has formalized the reimbursement decision-making process. Its transparency is the extent to which the arguments and motivation for decisions are clearly communicated to all stakeholders, including the public. Consistency refers to the extent to which repeated decisions have a similar process and outcomes. The validity of a decision is the extent to which the appraisal committee recommendations reflect the priorities within the society. All these contribute to accountability, which is the extent to which the organizational body can justify the decisions they take.

Our experiences in Europe have shown that explicitly stating, defining, and operationalizing the criteria that are used for assessment contributes to both the structure and transparency of the process. In Lombardy, the EVIDEM framework was adapted to local processes and decision procedures and benefits or shortfalls of several medical devices and procedures were identified and discussed through a performance-scoring exercise in 2 steps (personal then group discussion) against each criterion.⁴ In this case, having a pre-emptive list of explicit criteria to assist the performance assessment was a great improvement to the previously unstructured deliberative decision process. By doing so, arguments and motivations for decisions also can be clearly communicated to all stakeholders, including the public.

In the Netherlands, an example of how MCDA can increase structure and transparency of a decision process was the absence of an explicit comparator >

for performance assessment. It was identified in the project “A roadmap for uncertainty analysis in MCDA.”⁵ While evidence gathering and assessment in the Netherlands includes explicit comparison of costs and effects for the innovation compared to “current care,” for the other criteria, evidence on the relative performance is not always available. For instance, the budget impact of implementation (or de-implementation of current care) is difficult to estimate, and performance estimates had to be based on the committee members’ expert judgements. Lack of evidence to support decisions is recognized in the current decision-making process and reflected in the qualitative recommendations of the National Health Care Institute in the Netherlands. Criteria weighting and performance scoring in an MCDA model are complicated by the lack of a comparator and are likely reflected in higher uncertainty in weights and performance valuations of committee members. In MCDA, the impact of being uncertain about some of the inputs on the output of the model can be estimated in the final step of MCDA, and the validity of the decision can be appraised.⁶

If MCDA were to be implemented, it should be integrated with strong deliberative components to combine the benefits of the quantitative analysis with the benefits of a strong, value-based deliberation process, thus resulting in a sum that is better than its parts.

Unstructured qualitative decision processes can result in disproportionate time being spent on minor issues or undue attention to the opinions of more dominant panel members. When MCDA is used, the relative priorities and values of all committee members are elicited. In our opinion, this can focus the discussion on the most important issues and the most divergent opinions in the panel. In the discussion, each of the committee’s members is required to explicitly state the reasons and arguments to support their judgments, which benefits transparency and validity of the decision process (“Are our explicit priorities in line with how we feel about the importance of this issue, if not, why is this the case?”). Once the overall value of the alternative compared to current care is calculated via formal MCDA, it is of paramount importance to discuss the reimbursement recommendation deliberatively. If there is a feeling of unease with the proposed decision, arguments to deviate from the proposed decision are probably not part of the core set of criteria, and should be discussed (“What are the reasons, not considered in the analysis so far, that would influence our judgment? Would they increase or decrease the value of the innovation that is being discussed?”). When applied and well documented, this can bring validity (and in a broader perspective: accountability) to the decision process.

One would expect that by improving the structure, transparency, and validity of the decision process, the decisions that follow become more consistent. However, varying priorities between assessments reduce consistency of subsequent decisions. While these are difficult to recognize in qualitative discussion, quantitative weight elicitation can highlight inconsistency in importance of criteria. Once potential inconsistency is identified,

it can be a topic of discussion. The most extreme manner in which to increase consistency of repeated decisions within an MCDA framework is to use a fixed criteria set and equal criteria weights over decisions — performance scales that would be able to capture the value of a wide range of innovations and a direct link between overall value of an innovation and the decision to reimburse. For instance, by introducing cut-off points above which reimbursement of health innovations is recommended. However, the major pitfall of this approach is that it could reduce HTA to an algorithmic approach, and the specific aim of having decision committees, which is to appraise the evidence from a public perspective and incorporating societal values, would be lost.⁷

POTENTIAL PITFALLS TO USING MCDA TO SUPPORT HTA

Our experiences in Europe taught us that building a valid MCDA model to assist reimbursement decision making is a difficult but worthwhile process.^{4,5} First, it is questionable whether one MCDA decision model will fit all decisions made by HTA organizations. Some criteria (like disease severity) might not be relevant to consider if multiple innovations for one disease are considered, but are very important for prioritization of reimbursement on a higher level⁸. Second, preferential independence of criteria, which is one of the requirements for MCDA, is difficult to ensure in the model, as many commonly considered criteria are related in their current definition. For instance, cost-effectiveness, budget impact, and effectiveness are related. Third, if MCDA is used to support HTA, a balance has to be struck between explicitly formalizing all conceivable criteria, and keeping the set manageable for the larger organisation, for instance, with regard to evidence gathering and assessment. In addition, the explicit technology performance scoring and weighting of criteria in themselves add an additional layer of complexity to the decision process. The time requirements of doing so have to be considered. Fourth, for some criteria, a performance scale is not easily defined and developing scales takes much time. For instance, in recent years, the GRADE methodology was developed to measure the overall quality of clinical evidence [9]. However, quality of the evidence itself is influenced by multiple criteria, for instance, the number of studies available, their research methodology, the number of respondents per study, and of course, their findings. Further study is required to determine whether using GRADE to assess evidence will suffice or whether meta-analysis or more complicated modelling techniques are required to determine the quality of the clinical evidence as a whole. Fifth, with regard to data aggregation, a simple additive value function may not be able to capture non-compensatory criteria, which are employed by agencies, and more complex analysis methods might be required. Although this is technically possible within MCDA, models that are more complex are more difficult to understand by laypeople, who are the main audience to which HTA decisions have to be justified. Finally, and most importantly, although having an explicit list of criteria is important, we do not expect that all arguments in favor or against reimbursement during appraisal can be fully captured in a “one-size-fits-all” set of criteria. We envisage such a set, along with criteria weight elicitation and performance valuation as a starting point for appraisal, not the result.

When introducing MCDA to a European HTA decision process, a sensitive balance has to be struck between having the benefit of

increased structure and transparency by having clear steps and questions in the decision process, while maintaining the high quality, in-depth discussion on societal priorities and the ability to deviate from criteria as needed and based on well-motivated argumentations. If the latter were lost, the benefit of having an appraisal panel would be lost. Reimbursement decision making is an ethical problem for which the goal should not be to provide a mathematical solution. However, MCDA can be used as a way of more systematic thinking about reimbursement decisions and thereby fulfill its promise of adding transparency and validity to the current process. If MCDA were to be implemented, it should be integrated with strong deliberative components to combine the benefits of the quantitative analysis with the benefits of a strong, value-based deliberation process, thus resulting in a sum that is better than its parts.⁷ •

ADDITIONAL INFORMATION

This preceding article is based on an issues panel given at the ISPOR 19th Annual European Congress.

To learn more about the ISPOR MCDA Task Force, go to <https://www.ispor.org/Multi-Criteria-Decision-Analysis-guideline.asp>

REFERENCES

1. Belton V, Stewart T. Multiple Criteria Decision Analysis. Boston Dordrecht London: Kluwer Academic Publishers, 2002.
2. 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.
3. Marsh K, M IJ, 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: 125-137.
4. Castro H, Tringali M, Cleemput I, et al. Advancing MCDA and HTA into Coverage Decision-Making. In: Marsh K, Goetghebeur M, Thokala P, et al., (eds.). *Multi-Criteria Decision Analysis to Support Healthcare Decisions*. Cham: Springer International Publishing, 2017.
5. Development WTNOfHRa. Project: A roadmap for uncertainty analysis in MCDA.
6. Broekhuizen H, Groothuis-Oudshoorn CGM, van Til JA, et al. A Review and Classification of Approaches for Dealing with Uncertainty in Multi-Criteria Decision Analysis for Healthcare Decisions. *Pharmacoeconomics*. 2015; 33: 445-455.
7. Baltussen R, Paul Maria Jansen M, Bijlmakers L, et al. Value Assessment Frameworks for HTA Agencies: The Organization of Evidence-Informed Deliberative Processes. *Value Health* 2017;20:256-260.
8. Schmitz S, McCullagh L, Adams R, et al. Identifying and Revealing the Importance of Decision-Making Criteria for Health Technology Assessment: A Retrospective Analysis of Reimbursement Recommendations in Ireland. *Pharmacoeconomics* 2016;34:925-937.
9. Balshem H, Helfand M, Schunemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol* 2011;64:401-406.