Eur J Health Econon
DOI 10.1007/s10198-013-0517-9

ORIGINAL PAPER

A multi-criteria decision analysis perspective on the health economic evaluation of medical interventions

Douwe Postmus · Tommi Tervonen · Gert van Valkenhoef · Hans L. Hillege · Erik Buskens

Received: 6 February 2013 / Accepted: 26 June 2013

Abstract A standard practice in health economic evaluation is to monetize health effects by assuming a certain societal willingness-to-pay per unit of health gain. Although the resulting net monetary benefit (NMB) is easy to compute, the use of a single willingness-to-pay threshold assumes expressibility of the health effects on a single non-monetary scale. To relax this assumption, this article proves that the NMB framework is a special case of the more general stochastic multi-criteria acceptability analysis (SMAA) method. Specifically, as SMAA does not restrict the number of criteria to two and also does not require the marginal rates of substitution to be constant, there are problem instances for which the use of this more general method may result in a better understanding of the trade-offs underlying the reimbursement decision-making problem. This is illustrated by applying both methods in a case study related to infertility treatment.

Keywords Health economic evaluation · Multi-criteria decision analysis · Infertility treatment

D. Postmus (✉) · G. van Valkenhoef · H. L. Hillege · E. Buskens
Department of Epidemiology, University of Groningen, University Medical Center Groningen, P.O. Box 30.001, 9700 RB Groningen, The Netherlands
e-mail: d.postmus@umcg.nl

T. Tervonen
Econometric Institute, Erasmus University Rotterdam, Rotterdam, The Netherlands

H. L. Hillege
Department of Cardiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

JEL Classification C44 · I18

Introduction

In health economic evaluation, two or more alternatives (e.g., medical devices, drug therapies, or surgical procedures) are evaluated in terms of their costs and health consequences to assist health policy makers in deciding which of these alternatives should be reimbursed. Such analyses can be conducted using different methods and theoretical perspectives. In particular, a distinction is often made between cost-benefit analysis, where the objective is to ascertain the total amount that individuals would be willing to pay for the health effects resulting from a treatment (welfarist perspective) and cost-effectiveness analysis (sometimes referred to as cost-utility analysis when health effects are quantified in terms of quality-adjusted life years), where the objective is to maximize health effects given a budget constraint on the total health-care expenditure (resource allocation perspective) [4]. Although attempts to ascertain individuals' valuation of health effects are sometimes undertaken, the vast majority of published health economic evaluations are cost-effectiveness analyses [10]. In addition, given that cost-benefit analysis and cost-effectiveness analysis have fundamentally different philosophical underpinnings that cannot be reconciled in a meaningful way [7], it seems that the resource allocation perspective predominates in applied health economic analysis.

From a resource allocation perspective, the purpose of the health-care system is to maximize a health-related objective function subject to a budget constraint set by policy makers. However, as resource allocation decisions have to be made across a whole range of disease areas, it is

Published online: 11 July 2013

practically impossible to simultaneously compare the costs and effects of all possible treatments within each of these disease areas. The approach taken in the literature has therefore been to derive simple rules for reimbursement decision making from abstract mathematical formulations of the problem [16]. In particular, it has been shown that under some simplifying conditions, the optimal solution is to undertake all interventions whose ratio of costs to effects is below or equal to the maximum spendable amount per unit increase in the chosen measure of effectiveness, say γ, whose value depends on the size of the health-care budget [21]. From this, it follows that when choosing from a set of alternative treatments, it is optimal to first eliminate all (extendedly) dominated treatments and to then choose from the remaining treatments the one with the highest incremental cost-effectiveness ratio not exceeding γ [3, 8].

Although decision rules such as the one described in the previous paragraph can assist health policy makers in deciding whether a certain treatment should be eligible for reimbursement, a decision to reimburse this treatment should still be followed by a reallocation of the remaining budget as the inclusion of this treatment will result in a different value of γ (e.g., γ will increase if the new treatment replaces a more expensive treatment and decrease if relative to the current treatment the new treatment is both more effective and more costly). In practice, however, the dependence between γ and the available budget is often neglected [1]. Instead, it is usually assumed that γ is fixed at a certain value λ, which is generally referred to as the societal willingness-to-pay per unit of health gain. This reduces a rather complex resource allocation problem to a relatively straightforward unconstrained optimization problem, where the optimal alternative is simply the one that yields the highest net monetary benefit (NMB) [6].

Because of its simplicity, the NMB framework has become the current standard of practice in applied health economic evaluation. However, as the willingness-to-pay threshold is treated as an exogenous parameter, the use of this framework is no longer congruent with the resource allocation perspective from which it originated. In this article, we show that the practical application of the NMB framework nevertheless still has a theoretical foundation when the NMB function is interpreted as a value function as defined in multi-attribute value theory (MAVT). In particular, using this interpretation of the NMB function, we prove that the NMB framework can be seen as a special case of the more general stochastic multi-criteria acceptability analysis (SMAA) method, which, in the context of health policy decision making, has previously been applied to perform quantitative drug benefit-risk analysis of alternative treatments using all available evidence from a clinical trial [17] or a network of clinical trials [19]. In addition, as interpreting the NMB function as a value function implies that the reimbursement decision-making problem is treated as a multi-criteria decision problem, it is no longer required to restrict the number of criteria that are taken into account in a health economic evaluation to two. In fact, there are problem instances for which including more than two criteria results in a better understanding of the trade-offs underlying the reimbursement decision-making problem, as will be illustrated in a case study related to infertility treatment.

Multi-criteria decision analysis

Multi-criteria decision problems consist of a set of m alternatives that are evaluated in terms of n criteria. The vector of criteria measurements corresponding to alternative i is denoted by $x^i = (x^i_1, \ldots, x^i_n)$, where x^i_k represents the performance of alternative i on criterion k.

MAVT

In MAVT, a decision maker's preference structure is represented by a value function $v(x)$ such that alternative i is preferred over alternative j if and only if $v(x^i) > v(x^j)$. To simplify the assessment of $v(x)$, it is generally assumed that the criteria are preferentially independent [14]. The overall value function can then be written as a weighted additive combination of n partial value functions:

$$v(x, w) = \sum_{k=1}^{n} w_k \cdot v_k(x_k). \tag{1}$$

MAVT is based on the premise that a decrease in the performance on one criterion can be compensated by an increase in the performance on the other criteria. The partial value functions, normalized so that the worst possible level of performance on each criterion is assigned a value of 0 and the best possible level of performance is assigned a value of 1, are therefore constructed in such a way that equal decrements in v_k represent the same loss of value to the decision maker (and can thus be compensated by the same amount of increase in the performance on the other criteria). The weights w_k, normalized so that they sum to one, indicate how much more important the swing from worst to best on one criterion is compared to the swings from worst to best on the other criteria. For example, $w_k > w_l$ implies that if the decision maker had to choose between improving either criterion k or criterion l from the worst to the best value, he or she would improve criterion k. To emphasize that the weights in MAVT reflect the relative importance of swings between reference points on the criteria scales rather than the intrinsic importance of these criteria, they are sometimes referred to as swing weights [2]. All weights considered in this article should be interpreted as swing weights.

To use the additive value function for decision-support purposes, the partial value functions need to be defined and the weights need to be specified. However, given that many countries that have adopted the NMB framework are not willing to make explicit statements about their willingness-to-pay thresholds [9], this problem is also likely to occur for the weights if MAVT is applied to support reimbursement decision making. In addition, in MAVT, the values of x_k^i are considered to be deterministic. However, the criteria measurements used to inform reimbursement decision making are derived from clinical trials and/or observational studies (either directly or indirectly as parameter estimates of a mathematical model), and as such they are inherently uncertain. In such situations, SMAA [15] can be applied to determine (1) the 'typical' values of the weights that would make each alternative the preferred one and (2) the probability that an alternative obtains a certain rank given the uncertainty in the criteria measurements and/or the values of the weights.

SMAA

In SMAA, the weight vector $w = (w_1, \ldots, w_n)$ of the additive value function is assumed to be uniformly distributed in the feasible weight space W defined through a set of linear weight constraints, and the criteria measurements $\xi = (x^1, \ldots, x^m)$ are assumed to be random variables with joint density function $f_\Xi(\xi)$. For given realizations w of W and ξ of Ξ, the rank of each alternative is defined as an integer from the best rank (=1) to the worst rank (=m) by means of the ranking function

$$\mathrm{rank}(i, \xi, w) = 1 + \sum_{k=1}^{m} I(v(x^k, w) > v(x^i, w)), \tag{2}$$

where $I()$ is an indicator function returning the value 1 for I(true) and the value 0 for I(false). The *rank acceptability index* $b(i, r)$, defined as the probability that alternative i is positioned at rank r, can then be computed as

$$b(i, r) = \int_{w \in W} f_W(w) \int_{\xi \in \Xi : rank(i,\xi,w)=r} f_\Xi(\xi) d\xi dw\,. \tag{3}$$

In addition, the SMAA methods allow computing the weights a 'typical' decision maker supporting each alternative might have. These so-called *central weight vectors* can be presented to the decision maker to help him or her understand what kind of weights would favor a certain alternative, without providing preference information. More formally, the central weight vector of alternative i, denoted by w_c^i, is defined as the expected center of gravity of all possible weight vectors that position this alternative at rank 1:

$$w_c^i = \frac{1}{b(i,1)} \int_{w \in W} f_W(w) w \int_{\xi \in \Xi : rank(i,\xi,w)=1} f_\Xi(\xi) d\xi dw\,. \tag{4}$$

Finally, for a given weight vector w, the *confidence factor* $p(i, w)$, defined as the probability that alternative i to positioned at rank 1 when w is used to represent the decision maker's preferences, can be computed as

$$p(i, w) = \int_{\xi \in \Xi : rank(i,\xi,w)=1} f_\Xi(\xi) d\xi. \tag{5}$$

When computed for the central weight vectors, the confidence factors $p(i, w_c^1), \ldots, p(i, w_c^m)$ indicate whether the criteria measurements are sufficiently accurate to discern the efficient alternatives (i.e., all alternatives with a non-zero first rank acceptability index). Central weight vectors with low confidence factors (<0.50) should be interpreted with care: even if a decision maker finds a central weight vector to correspond to his or her preferences, there might be another alternative that achieves a higher first rank acceptability with the given weights.

The NMB framework

Consider m alternatives that are evaluated in terms of their cost c^i and effectiveness e^i. The NMB of alternative i is defined as

$$\mathrm{NMB}(c^i, e^i, \lambda) = \lambda e^i - c^i, \tag{6}$$

where λ denotes the willingness-to-pay per unit increase in the chosen measure of effectiveness. The best alternative is the one that yields the highest NMB.

To take into account uncertainty in the criteria measurements, it is usually assumed that $\xi = ((c^1, e^1), \ldots, (c^m, e^m))$ is a random vector with joint density function $f_\Xi(\xi)$. The probability $P(i, \lambda)$ that alternative i has the highest NMB can then be computed as

$$P(i, \lambda) = \int_{\xi \in \Xi : \mathrm{RANK}(i,\xi,\lambda)=1} f_\Xi(\xi) d\xi, \tag{7}$$

where $\mathrm{RANK}(i, \xi, \lambda) = 1 + \sum_{k=1}^{m} I\ [\mathrm{NMB}(c^k, e^k, \lambda) > \mathrm{NMB}(c^i, e^i, \lambda)]$. As the exact value of λ is generally unknown, decision support is typically provided by plotting $P(i, \lambda)$ against λ, resulting in the alternative's cost-effectiveness acceptability curve (CEAC) [20].

The NMB framework as a special case of SMAA

For a given value of λ, the NMB function uniquely specifies the decision maker's preference structure. In the nomenclature of MAVT, the NMB function therefore is a value function. However, this representation of the decision maker's preference structure is not unique: any strictly monotone transformation of the NMB function results in the same preferential ordering of any set of points in the cost-effectiveness space and therefore represents the same preference structure [14]. For example, the functions $\mathrm{NMB}(c^i, e^i, \lambda)$, $\log(\mathrm{NMB}\ c^i, e^i, \lambda))$, and $\sqrt{\mathrm{NMB}(c^i, e^i, \lambda)}$ all result in the same ranking of the decision alternatives and must therefore represent the same preference structure.

To prove that the NMB framework can be seen as a special case of the more general SMAA method, we need to show that (1) there exists a weighted additive function that represents the same preference structure as the NMB function and (2) that when using this function the information presented in a CEAC can be derived from the SMAA descriptive indices. Both the NMB function and a weighted additive combination of two linear partial value functions exhibit constant marginal rates of substitution (i.e., the amount that the decision maker is willing to pay for a unit increase in health effects depends on neither c nor e). Intuitively, it must therefore be possible to construct a weighted additive combination of two linear partial value functions that represents the same preference structure as the NMB function. To prove this formally, consider the linear partial value functions $v_1(e) = \frac{e-\underline{e}}{\overline{e}-\underline{e}}$ and $v_2(c) = \frac{\underline{c}-c}{\underline{c}-\overline{c}}$, where $\underline{c}$ and $\overline{c}$ denote the worst and best possible value of the cost criterion, and $\underline{e}$ and $\overline{e}$ denote the worst and best possible value of the effectiveness criterion. Then, it follows that the NMB function can be expressed as

$$\mathrm{NMB}(c,e,\lambda) = (\lambda\overline{e} - \lambda\underline{e} + \underline{c} - \overline{c})(w_1 v_1(e) + w_2 v_2(c)) + \lambda\underline{e} - \underline{c}, \tag{8}$$

with w_1 and w_2 defined as

$$w_1 = \frac{(\lambda\overline{e} - \lambda\underline{e})}{\lambda\overline{e} - \lambda\underline{e} + \underline{c} - \overline{c}}, \tag{9}$$

$$w_2 = \frac{\underline{c} - \overline{c}}{\lambda\overline{e} - \lambda\underline{e} + \underline{c} - \overline{c}}. \tag{10}$$

Hence, the NMB function can be written as a positive affine transformation of the weighted additive function $v((c, e), w) = w_1 v_1(c) + w_2 v_2(e)$, thereby implying that these two functions represent the same preference structure. The details of this derivation are provided in the "Appendix". In addition, by substituting (9) and (10) in (5), it follows that $P(i, \lambda)$ is simply the confidence factor $p(i, w)$ corresponding to the weight vector $w = \left(\frac{(\lambda\overline{e}-\lambda\underline{e})}{\lambda\overline{e}-\lambda\underline{e}+\underline{c}-\overline{c}}, \frac{\underline{c}-\overline{c}}{\lambda\overline{e}-\lambda\underline{e}+\underline{c}-\overline{c}}\right)$ of the additive value function $v((c, e), w)$. CEACs can therefore be seen as graphical presentations of the SMAA confidence factors for the two-criteria problem.

To conclude, both the NMB framework and SMAA are based on the assumption that the decision maker's preference structure can be represented by an additive value function. However, unlike the NMB framework, SMAA does not restrict the number of criteria to two, and, although assumed in the case study presented in the next section, it does not require the partial value functions to be linear. In addition, as both methods combine an additive value function with stochastic simulation to derive similar descriptive indices supporting the decision problem, the NMB framework can be seen as a special case of the more general SMAA method.

Application to infertility treatment selection

To illustrate the practical usefulness of SMAA for health economic evaluation, we employed a previously published mathematical model for infertility treatment [12]. The objective of the original study was to compare the cost-effectiveness of seven in-vitro fertilization (IVF) strategies based on a maximum of three consecutive IVF cycles and different combinations of the following embryo transfer policies per cycle: *elective single embryo transfer* (eSET), *double embryo transfer (DET)*, and *standard treatment practice* (STP), consisting of eSET in patients <38 years of age and DET in the remainder of patients. For this purpose, an elaborate Markov cycle tree was developed, taking into account canceled cycles, availability of only one embryo (resulting in a so-called compulsory single embryo transfer), declining pregnancy rates in subsequent cycles, the possibility of having frozen embryo transfers when a patient did not achieve a pregnancy after fresh embryo transfer or had a miscarriage/stillborn child, and treatment dropouts due to cycle cancellation or fertilization failure. The cycle time used in the model was defined as one IVF cycle, and the time horizon was defined as a maximum of three cycles in all strategies. For full details of the model structure, the seven IVF strategies, and the probability distributions of the model parameters, the reader is referred to the original publication [12].

Criteria

The different IVF strategies were evaluated on three criteria: cost, treatment success, and adverse consequences. Treatment success was quantified in terms of the mean live birth probability for a couple starting IVF treatment, whereas the adverse consequences were expressed in terms of the risk of a twin pregnancy. Costs were analyzed from a societal perspective and included the cost of the IVF treatment, the cost of a singleton and twin pregnancy, the cost of delivery, and the cost of the period from birth until 6 weeks after birth.

Joint probability distribution of the criteria measurements

The joint probability distribution of the criteria measurements was approximated numerically by repeatedly simulating from the Markov cycle tree for different values of the model parameters. Table 1 summarizes the results of 10,000 Monte Carlo iterations (in each iteration, 5,000 runs of the Markov cycle tree were performed to obtain stable estimates for the current realizations of the model parameters). The same seed values for generating these parameters were applied to each treatment strategy, resulting in a non-zero correlation among the criteria measurements of the seven IVF strategies.

NMB analysis

We performed an NMB analysis with the probability of a live birth as the selected measure of effectiveness and the value of λ assumed to be unknown but contained within the interval [0, 80k]. Irrespective of the value of λ, the four hybrid IVF strategies always had very low (<0.15) probabilities of yielding the highest NMB. It therefore seems reasonable to eliminate these four strategies from our inquiry. The CEACs of the three remaining strategies are provided in Fig. 1. These curves show that the optimal strategy strongly depends on the value of λ, with 3× eSET being optimal for low values of λ, 3× STP being optimal for intermediate values of λ, and 3× DET being optimal for high values of λ. Without any additional information on the value of λ, it is therefore impossible to further reduce the set of acceptable IVF strategies. Hence, by applying the NMB framework, we were able to eliminate four IVF strategies, but it remains difficult to choose among the three remaining strategies.

Table 1 Median (interquartile range) of the criteria measurements

Strategy	Live birth	Twin pregnancy	Cost
3× eSET	0.37 (0.06)	0.008 (0.005)	14,294 (2,868)
eSET + 2× STP	0.46 (0.05)	0.015 (0.005)	15,290 (2,922)
eSET + STP + DET	0.47 (0.04)	0.031 (0.010)	15,764 (2,974)
eSET + 2× DET	0.49 (0.04)	0.063 (0.021)	16,569 (3,138)
3× STP	0.52 (0.06)	0.021 (0.006)	15,631 (2,961)
STP + 2× DET	0.55 (0.04)	0.063 (0.022)	16,749 (3,092)
3× DET	0.58 (0.05)	0.115 (0.039)	17,837 (3,462)

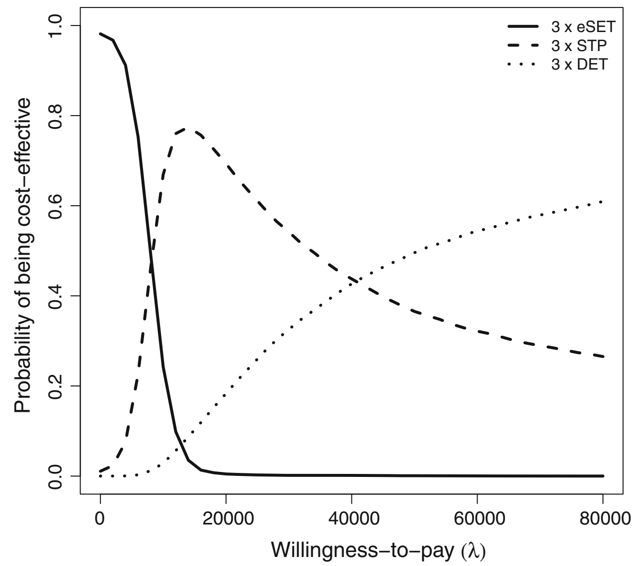

Fig. 1 Results of the NMB analysis

SMAA analysis

We performed an SMAA analysis with the criteria measurements linearly rescaled to the interval [0, 1] using the worst and best possible values as provided in Table 2 and the weight vector w assumed to be uniformly distributed over the two-simplex $\left\{w \in R^3 : \sum_{k=1}^{3} w_k = 1,\ w \geq 0\right\}$. The rank acceptability indices resulting from this analysis are depicted in Fig. 2, and the central weights and corresponding confidence factors are provided in Table 3. The hybrid IVF strategies both have low (<0.10) first rank acceptabilities and central weight vectors with low (<0.20) confidence factors. In line with the analysis based on the NMB framework, it is therefore still unlikely that switching between embryo transfer policies in consecutive IVF cycles is optimal. For the three remaining strategies, the rank acceptability indices reveal that the optimality of 3× eSET and 3× DET strongly depends on the decision maker's preferences (i.e., both strategies still have relatively low acceptabilities for the best rank but rather high acceptabilities for the worst rank). By looking at the central weights, we can learn more about the preferences that would favor these alternatives. In particular, it follows that 3× eSET is likely to be optimal when controlling cost and reducing the risk of a twin pregnancy is considered to be

Table 2 Scale ranges of the partial value functions

Criterion	Preference direction	Worst value	Best value
Live birth	↑	0.21	0.71
Twin pregnancy	↓	0.249	0.001
Cost	↓	32,944	8,245

far more important than improving the probability of treatment success, whereas 3× DET could be a viable strategy when the probability of a live birth is by far the most important criterion. In contrast, 3× STP is a compromise strategy that performs well in virtually all other settings. In the absence of preference information from the decision maker, this strategy therefore seems most suitable to adopt in daily clinical practice.

Discussion

In this article, we have proven that the NMB framework is a special case of the more general SMAA method, which does not restrict the number of criteria to two and does not require the partial value functions to be linear. Although the NMB framework is easy to understand and implement, it also forces analysts to select a single effectiveness criterion. For preventive and curative interventions (e.g., prevention of diabetes, cancer treatment), this can usually be effectively dealt with by aggregating all relevant health effects into an overall measure of effectiveness, such as (quality-adjusted) life years. However, as we have seen in our case study, the health consequences of interventions that are not directly targeted at prolonging a subject's life years are often much harder to capture in terms of a single measure of effectiveness. Applying the NMB framework to evaluate these latter interventions may lead to sub-optimal reimbursement decisions as the need to arbitrarily select one of the available outcome measures as the effectiveness criterion implies that only part of the health effects are accounted for in the decision-making process. In addition, the cost criterion is generally composed of several attributes, such as resource use within the health sector, out-of-pocket expenses by patients and their families, and productivity losses [8]. Because each of these costs are borne by different stakeholders, health policy makers may want to value them differently from each other. However, as the NMB framework is based on a single willingness-to-pay threshold, this is not possible. SMAA, in contrast, can be

Table 3 Central weights and corresponding confidence factors (CF)

Strategy	CF	Live birth	Twin pregnancy	Cost
3× eSET	0.801	0.066	0.527	0.407
eSET + 2× STP	0.115	0.224	0.274	0.502
eSET + STP + DET	0.005	0.399	0.245	0.356
eSET + 2× DET	0.002	0.580	0.151	0.269
3× STP	0.870	0.342	0.304	0.354
STP + 2× DET	0.189	0.574	0.249	0.177
3× DET	0.474	0.687	0.216	0.097

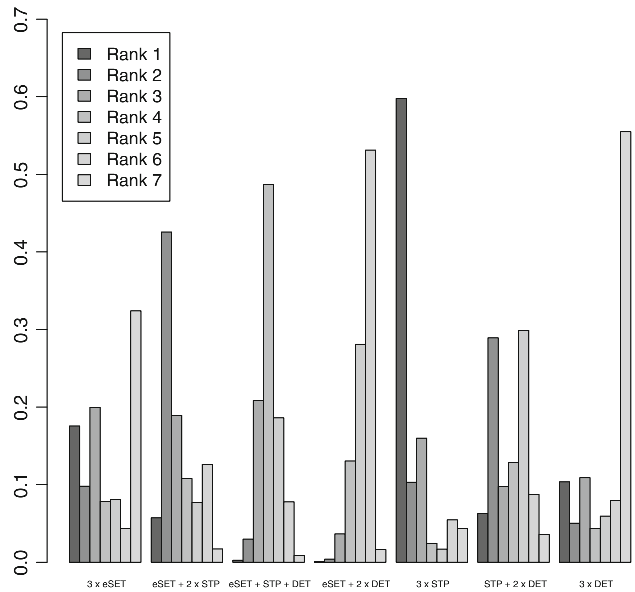

Fig. 2 Rank acceptability indices

applied with an arbitrary number of criteria, meaning that the above problems no longer apply when this approach is selected to support reimbursement decision making. The price that has to be paid for applying this more general method is that the weights no longer have the intuitive meaning of willingness-to-pay thresholds but should rather be interpreted as scaling factors that make unit increases in the partial value functions commensurate [5]. This implies that some education of the policy makers may be required before SMAA can successfully be applied to real-life reimbursement decision-making problems.

Within the NMB framework, CEACs offer a useful way of representing the decision uncertainty when there are two alternatives, say an intervention and a control. In such situations, the CEAC for the control is simply the complement of the CEAC for the intervention, meaning that only the CEAC for the intervention has to be provided. When more than two alternatives are considered, multiple CEACs have to be presented (i.e., one for each alternative). The standard way of visualizing these CEACs is by plotting them in the same figure, so that the curves sum to a probability of one vertically [11]. While such a representation depicts the probability of making the correct decision when a certain alternative is selected (e.g., for reimbursement of implementation), it provides no information on the alternative's probability distribution over the other ranks when making a wrong decision. This may make it difficult to identify good compromise solutions as it causes extreme alternatives (i.e., alternatives with large probabilities of being ranked at either the first or the last place and small probabilities of being ranked at any of the intermediate places) to look similar to those with large probabilities for the best ranks only. One way to avoid this problem would be to plot not only the probability that an alternative is ranked best but also the probabilities that this alternative is ranked worst or ranked at any position between best and worst. Hence, even for problem instances for which the use of the NMB framework is considered to be appropriate, the SMAA rank acceptability indices could still be of value in summarizing the decision uncertainty (e.g., by computing these indices conditional on specific values of the willingness-to-pay threshold).

The application of multi-criteria decision analysis to reimbursement decision making has been proposed previously [13, 18], but not, as we aimed for in this article, with the objective to generalize the way in which health economic evaluations are currently conducted. Instead, these papers focused on augmenting the results of a cost-effectiveness analysis with other factors that may influence reimbursement decision making but are generally not included in a formal health economic evaluation, such as patient compliance, quality of evidence, and budget impact. We acknowledge that epidemiological, social, and ethical factors may play an important role in the decision-making process and that it therefore makes sense to consider them in a decision analysis. However, we would also like to point out that not all of these factors are equally suitable to be included as criteria in a supporting multi-criteria model. For example, although factors related to the relevance and quality of the data impact the precision with which the criteria are measured, they are not features of treatments among which trade-offs can be established. Such factors should therefore not be included as criteria in a supporting multi-criteria model but should rather be used to guide the selection of the input parameters for which a subsequent sensitivity analysis is worthwhile to conduct. Also, budget impact is obviously an important factor in the reimbursement decision-making process. However, rather than including it as a criterion in a supporting multi-criteria model, it may be more fruitful to consider budget impact when establishing the relative importance of the included cost and effectiveness criteria since it seems likely that decision makers are willing to pay less for a certain average increase in health effects as the budget impact increases.

Acknowledgments This research was conducted within the framework of the Center for Translational Molecular Medicine, project PREDICCt (grant 01C-104) and supported by the Dutch Heart Foundation, Dutch Diabetes Research Foundation, and Dutch Kidney Foundation.

Appendix

Define $V_1(e) = \lambda\ e$ and $V_2(c) = -c$, and consider the NMB function

$$\mathrm{NMB}(c^i, e^i, \lambda) = \lambda e^i - c^i = V_1(e^i) + V_2(c^i). \qquad (11)$$

Let $\underline{c}$ and $\overline{c}$ denote the worst and best possible value of the cost criterion, and let $\underline{e}$ and $\overline{e}$ denote the worst and best possible value of the effectiveness criterion. This allows us to express $V_1(e)$ and $V_2(c)$ as the following positive affine transformations of the linear partial value functions $v_1(e) = \frac{e-\underline{e}}{\overline{e}-\underline{e}}$ and $v_2(c) = \frac{c-\underline{c}}{\overline{c}-\underline{c}}$:

$$V_1(e) = \frac{\lambda\overline{e} - \lambda\underline{e}}{\lambda\overline{e} - \lambda\underline{e}}(\lambda e - \lambda\underline{e}) + \lambda\underline{e} = (\lambda\overline{e} - \lambda\underline{e})v_1(e) + \lambda\underline{e}, \qquad (12)$$

$$V_2(c) = \frac{\underline{c} - \overline{c}}{\underline{c} - \overline{c}}(\underline{c} - c) - \underline{c} = (\underline{c} - \overline{c})v_2(c) - \underline{c}. \qquad (13)$$

Now, by substituting (12) and (13) in (11), it follows after rewriting that

$$\mathrm{NMB}(c, e, \lambda) = (\lambda\overline{e} - \lambda\underline{e})v_1(e) + (\underline{c} - \overline{c})v_2(c) + \lambda\underline{e} - \underline{c}. \qquad (14)$$

Finally, by normalizing the scaling factors, we obtain the following expression for the NMB function:

$$\mathrm{NMB}(c, e, \lambda) = (\lambda\overline{e} - \lambda\underline{e} + \underline{c} - \overline{c})(w_1 v_1(e) + w_2 v_2(c)) + \lambda\underline{e} - \underline{c}, \qquad (15)$$

with w_1 and w_2 defined as

$$w_1 = \frac{(\lambda\overline{e} - \lambda\underline{e})}{\lambda\overline{e} - \lambda\underline{e} + \underline{c} - \overline{c}}, \qquad (16)$$

$$w_2 = \frac{\underline{c} - \overline{c}}{\lambda\overline{e} - \lambda\underline{e} + \underline{c} - \overline{c}}. \qquad (17)$$

References

1. Al, M., Feenstra, T., Van Hout, B.: Optimal allocation of resources over health care programmes: dealing with decreasing marginal utility and uncertainty. Health Econ. **14**(7), 655–667 (2005). doi:10.1002/hec.973
2. Belton, V., Stewart, T.J.: Multiple criteria decision analysis—an integrated approach. Kluwer Academic Publishers, Dordrecht (2002)
3. Briggs, A.H., Claxton, K., Sculpher, M.: Decision modelling for health economic evaluation. Oxford University Press, Oxford (2006)
4. Brouwer, W., Koopmanschap, M.: On the economic foundations of CEA. ladies and gentlemen, take your positions! J. Health Econ. **19**(4), 439–459 (2000). doi:10.1016/S0167-6296(99)00038-7
5. Choo, E.U., Schoner, B., Wedley, W.C.: Interpretation of criteria weights in multicriteria decision making. Comput. Ind. Eng. **37**(3), 527–541 (1999). doi:10.1016/S0360-8352(00)00019-X
6. Claxton, K., Posnett, J.: An economic approach to clinical trial design and research priority-setting. Health Econ. **5**(6), 513–524 (1998). doi:10.1002/(SICI)1099-1050(199611)5:6<513::AID-HEC237>3.0.CO;2-9

7. Dolan, P., Edlin, R.: Is it really possible to build a bridge between cost-benefit analysis and cost-effectiveness analysis? J. Health Econ. **21**(5), 827–43 (2002). doi:10.1016/S0167-6296(02)00011-5
8. Drummond, M., Sculpher, M., Torrance, G., O'Brien, B., Stoddart, G.: Methods for the economic evaluation of health care programmes. Oxford University Press, New York (2005)
9. Eichler, H., Kong, S., Gerth, W., Mavros, P., Jönsson, B.: Use of cost-effectiveness analysis in health-care resource allocation decision-making: how are cost-effectiveness thresholds expected to emerge? Value in Health **7**(5), 518–528 (2004). doi:10.1111/j.1524-4733.2004.75003.x
10. Elixhauser, A., Halpern, M., Schmier, J., Luce, B.: Health care CBA and CEA from 1991 to 1996: an updated bibliography. Med. Care **36**(5), MS1–MS9 (1998). doi:10.1097/00005650-199805001-00001
11. Fenwick, E., Claxton, K., Sculpher, M.: Representing uncertainty: the role of cost-effectiveness acceptability curves. Health Econ. **10**(8), 779–787 (2001). doi:10.1002/hec.635
12. Fiddelers, A.A.A., Dirksen, C.D., Dumoulin, J., van Montfoort, A., Land, J.A., Janssen, J.M., Evers, J.L.H., Severens, J.L.: Cost-effectiveness of seven IVF strategies: results of a Markov decision-analytic model. Hum. Reprod. **24**(7), 1648–1655 (2009). doi:10.1093/humrep/dep041
13. Goetghebeur, M., Wagner, M., Khoury, H., Levitt, R., Erickson, L., Rindress, D.: Bridging health technology assessment (HTA) and efficient health care decision making with multicriteria decision analysis (MCDA): applying the evidem framework to medicines appraisal. Med. Decis. Making **32**(2), 376–388 (2012). doi:10.1177/0272989X11416870
14. Keeney, R., Raiffa, H.: Decisions with multiple objectives: preferences and value tradeoffs. Wiley, New York (1976)
15. Lahdelma, R., Salminen, P.: SMAA-2: Stochastic multicriteria acceptability analysis for group decision making. Oper. Res. **49**(3), 444–454 (2001). doi:10.1287/opre.49.3.444.11220
16. Stinnett, A., Paltiel, A. et al.: Mathematical programming for the efficient allocation of health care resources. J. Health Econ. **15**(5), 641–654 (1996). doi:10.1016/S0167-6296(96)00493-6
17. Tervonen, T., van Valkenhoef, G., Buskens, E., Hillege, H.L., Postmus, D.: A stochastic multi-criteria model for evidence-based decision making in drug benefit-risk analysis. Stat. Med. **30**(12), 1419–1428 (2011). doi:10.1002/sim.4194
18. Thokala, P., Duenas, A.: Multiple criteria decision analysis for health technology assessment. Value Health **15**(8), 1172–1181 (2012). doi:10.1016/j.jval.2012.06.015
19. van Valkenhoef, G., Tervonen, T., Zhao, J., de Brock, B., Hillege, H.L., Postmus, D.: Multi-criteria benefit-risk assessment using network meta-analysis. J. Clin. Epidemiol. **65**(4), 394–403 (2012). doi:10.1016/j.jclinepi.2011.09.005
20. Van Hout, B., Al, M., Gordon, G., Rutten, F.: Costs, effects and c/e-ratios alongside a clinical trial. Health Econ. **3**(5), 309–319 (1994). doi:10.1002/hec.4730030505
21. Weinstein, M., Zeckhauser, R.: Critical ratios and efficient allocation. J. Public Econ. **2**(2), 147–157 (1973). doi:10.1016/0047-2727(73)90002-9