Approaches to Aggregation and Decision Making—A Health Economics Approach: An ISPOR Special Task Force Report

**Abstract**

The fifth paper of our Special Task Force (STF) report identifies and discusses two aggregation issues: (a) aggregation of cost and benefit information across individuals to a population level for benefit plan decision making, and (b) combining multiple elements of value into a single value metric for individuals. First, we argue that additional elements could be included in measures of value, but such elements have not generally been included in measures of QALYs. As an example, we describe a recently developed extended cost-effectiveness analysis (ECEA) that provides a good example of how to use a broader concept of utility. ECEA adds two features—measures of financial risk protection as well as income distributional consequences. We then discuss a further option for expanding this approach, -- "augmented" CEA that can introduce many of the value measures discussed in Section 2. Neither of these approaches, however, provide a comprehensive measure of value. To resolve this issue, we review techniques called multi-criteria decision analysis (MCDA) that can provide a comprehensive measure of value. We then discuss budget setting and prioritization using MCDA, issues not yet fully resolved. Next we discuss deliberative processes, which represent another important approach for population- or plan-level decisions used by many health technology assessment bodies. These utilize quantitative information on CEA and other elements, but the group decisions are reached by a deliberative, voting process. Finally, we briefly discuss use of stated preference methods for developing "hedonic" value frameworks, and conclude with some recommendations in this area.

## 5.1 Introduction

This section identifies and discusses two types of aggregation issues. One arises from the aggregation of cost and benefit information across individuals to a population level for benefit plan decision making. The other deals with the combination of multiple elements of value into a single value metric for individuals. We assess both issues here.

Regarding the first, CEA and (more broadly) cost-benefit analysis (CBA) share common origins and common deficits. Since they originate in utility maximization for a single individual, they cannot accommodate societal issues that involve (among other things) the distribution of costs and benefits across a population. For example, the US 1939 Flood Control Act specified that water projects could be undertaken only when "the benefits, *to whomever they accrue*, [be] in excess of the estimated costs." [*Our emphasis.*] These "distributional issues" remain outside the domain of traditional CEA and CBA. Thus, the problem is how best to aggregate the benefits enjoyed and the costs borne by individuals into a societal statement of value. Bator proposes (in effect) a benign dictator whose utility function includes the utilities of individuals in the population, allowing different weights for different people [1], but this is seldom if ever operationalized meaningfully.

Even the question of "distribution" can have multiple meanings. For example, regarding the health of different groups in a population, does "equity" mean equal access to health care or does it mean equal health outcomes? The latter approach implies a greater emphasis on caring for population subgroups with poor health status than does the former. These different meanings of "equity" have different implications for the way one might prioritize various health interventions.

The second class of aggregation issues appears when the definition of value has more than a single element or dimension. CEA and CBA each have a single metric of value—cost per QALY (or some similar measure of health benefit such as disability-adjusted life year (DALY) or others) or net monetary value (in CBA). Some people object to having the analyst monetize health benefits (as required by CBA), and hence prefer CEA. When multiple dimensions of value arise, CBA and CEA have no way to formally incorporate them (except, in CBA, by monetizing them, where feasible). Thus, this raises the issue of aggregation across value components.

In Section 2, we described a number of decision contexts and perspectives potentially relevant to considering and estimating value in a microeconomic framework [2]. The two types of aggregation issues are then related in that the aggregation of elements of value at the individual level becomes a necessary input for the aggregation to the population level. As mentioned, most CEAs—at least those assessing new biopharmaceutical products—have been oriented to the normative question of whether or not a health plan should adopt a new technology. Thus, they can be an important element in health technology assessment (HTA) which often considers a broader range of components or issues.

We discuss these issues below, organizing this discussion as follows. First, in Section 5.2, we argue that additional elements could be included in measures of value, but such elements have not generally been included in measures of QALYs. As a specific example, in Section 5.3 we describe a recently developed methodology called extended CEA (ECEA) that provides a good example of how to use a broader concept of utility, which has been applied to some specific interventions. ECEA adds two features—the element of financial risk protection (one part of insurance value assessed in Section 2) as well as income distributional consequences [2]. Section 5.4 discusses further options for expanding this approach (what we call "augmented" CEA). In Section 5.5 we discuss another, even broader approach called multi-criteria decision analysis (MCDA) that has some intuitive appeal, especially for decision makers who consider other elements or components not easily measured in CEA or CBA. This approach requires that decision makers place quantitative weights on the various criteria, depending on their relative importance for a decision. This exercise can make group decision making more transparent. Section 5.6 discusses budget setting and prioritization in these frameworks, building upon previous discussion in Section 4 discussing the situation where budget allocations and decision-making cutoffs are not perfectly aligned [3]. Section 5.7 discusses deliberative processes which represent another important approach for population- or plan-level decisions that is used by many HTA bodies. It utilizes quantitative information on CEA and other elements, but in the end, the group decision is reached by a deliberative, voting process. This section includes a brief discussion of alternative voting methods that deliberative bodies might use.

**5.2 Additional Value Elements: Conceptually Appealing but Currently Impractical to implement**

There remain a number of these other elements, some discussed in Section 3, that analysts *in concept* could include in a CEA (or CBA) structure, but cannot meaningfully do so because relevant data do not exist, and are impractical for real-world analysts at present [4]. Consider, as one example, a vaccine against or treatment of a virulent disease such as Zika, Ebola, or historical predecessors such as leprosy, tuberculosis, poliomyelitis, SARS, and AIDS. Public policy in these cases is often driven by issues such as public fear of contagion that are not readily captured in traditional CEA models. Such models normally do not include the entire at-risk population and their disutility from the threat of the disease. The missing element is the loss of utility associated with the potential risk of exposure to the disease—the "fear factor." This fear has dominated public policy towards these diseases for millennia, resulting in isolation colonies for people with leprosy, quarantine rules for tuberculosis, travel bans during SARS

epidemics, and many other similar actions. An associated issue is the utility gained from permanently eliminating the disease, raising a further important issue—viz. aggregating costs and benefits across future generations of the population.

An additional array of potential value elements fall into this general category of "conceptually feasible but generally impractical" for use in CEA or CBA without further research and testing. This includes, for example, such things as: fit with existing infrastructure/programs; availability of requisite trained personnel; improvement in children's school participation and highest-grade attainment (and hence future earnings); ethical considerations (e.g., involving end-of-life alternatives or manipulation of genetic material); and—for an individual patient—fears associated with specific types of therapies (e.g., radiation, genetic modification), as well as other elements discussed in Section 3. In what follows, we discuss methods that could be used to include both these less tractable and the more traditional value elements in a consolidated valuation.

**5.3 Extended Cost-Effectiveness Analysis (ECEA)**

In principle, one can conceive of estimating or eliciting from individuals the total willingness to pay for a defined bundle of attributes. In practice, this is often difficult to do because some key elements– e.g., the willingness to pay for scientific spillovers, for elimination of fear, for added clinical certainty, or for less inequality—remain unknown. Absent this perfect solution, some authors have proposed strategies for aggregating a few key elements of value into a systematic framework.

An approach that also captures important aspects of equity, albeit not within a fully aggregated value measure, is extended cost-effectiveness analysis (ECEA) [5]. ECEA aims to describe the value of medical interventions and policies along three specific dimensions: health gains, financial risk protection, and social cost, thus extending beyond the traditional CEA measures of health gains and costs. As a further refinement (separate from gathering information about these added dimensions of value), ECEA collects this information for different sub-populations of interest, thus allowing consideration of distribution of benefits and costs (equity and fairness). Financial risk protection reflects the reduction in the risk of financial distress due to healthcare costs, or the financial costs of poor health. It represents a subset of the "insurance value" concept discussed in Section 3 [4]. Cost is measured from the perspective of the society as a whole, as is common in CEA. One could envision including additional measures of value, but ECEA consists of these three. Equity reflects the value of promoting a more just distribution of access to health care or health outcomes themselves or conversely, the loss arising from expanding inequality.

ECEA evaluates the distributional effects of a medical intervention by specifying a set of population subgroups of interest. For instance, one might evaluate groups that vary by income, education, baseline health, diseases status, or some other characteristics of interest. By subgroup, ECEA then reports the health gains, private expenditures avoided, and financial risk protection. Health gains can be measured using QALYs or other suitable units. Private expenditures avoided reflect the reduction in out-of-pocket spending by subgroup. Theoretically, this includes the totality of effects on private spending, including changes in non-health spending. Financial risk protection is measured as expenditures avoided suitably scaled by an individual's income. The concept is that a given level of avoided expenditures has value in proportion to a person's or household's income or assets.

Concerning the concept of risk protection, to substitute for the conceptually desirable risk-premium,[1] a variety of metrics are possible, including: expenditures averted as a fraction of income; the change in the share of individuals above a certain threshold of healthcare spending as a fraction of income; change in the number of individuals with disposable income falling below poverty thresholds; the willingness to pay for the reduction in expenditures faced in a given health state, in the context of a specific consumer utility function; and so on. Much can be written about each of these approaches, but the common feature is some measure of expenditures avoided, relative to the individual's willingness to bear financial burden. None of these alternatives measure perfectly the aspect of risk that is captured by variance in financial outlays and embedded in the standard risk-aversion measure.

ECEA then reports these three measures of benefit by subgroup and then scaled by the total cost of the intervention. Thus, one might report total cases of poverty averted per dollar of total cost, or dollars of financial risk-protection per dollar of total cost. These measures are analogous to the more conventional cost-effectiveness ratio of health gains (e.g., QALYs gained) per dollar of cost, but there is no standard measure of willingness to pay for these equivalent to a cutoff value for cost per QALY or other CEA investment rule.

Further, ECEA provides little guidance on how to choose a single metric to combine these elements of value. Indeed, the absence of a single unifying economic framework for ECEA makes this a matter of judgment rather than analysis. While some may view the multi-dimensional nature of ECEA as advantageous, others may view it as a limitation, because ECEA offers no obvious way to compare dimensions such as financial risk protection and health gains. Nonetheless, ECEA adds an important tool to the literature on healthcare value, particularly for analysts who care about and wish to quantify issues relating to equity and risk, and it may provide a useful stepping-stone to more complete approaches that we discuss in Section 5.5.

### 5.4 "Augmented Cost-Effectiveness Analysis"

Going beyond the three dimensions of utility chosen in ECEA, one could consider adding still more measures of value in addition to health gains, social costs, and financial risk protection. To differentiate this from ECEA, we describe this approach as an "augmented" cost effectiveness analysis (ACEA) ,emphasizing the extra dimensions of utility captured beyond health gains, costs, and financial risk)*.* These added elements of value might include (as examples) the value of physical risk protection, scientific spillovers, the value of reducing diagnostic uncertainty for patients, option value, insurance value, the value of preventing incurable diseases (such as polio, Ebola, or Zika), or the value of completely eliminating a contagious disease from the earth (as with smallpox, and nearly so with polio).

There are several potential approaches to including these added elements. One is to create a separate set of element-by-element comparisons, perhaps by subgroup if relevant. However, this approach—as

[1] The purest measure from standard welfare economics analysis would be the Pratt-Arrow risk premium, i.e. the product of the absolute risk aversion measure (r) and 0.5 times the variance of remaining out of pocket expenditures (scaled to income), [6,7]. Both of which are difficult to measure or even approximate at the individual level Garber and Phelps reference the corporate finance literature to estimate r* = -r*Income between 1 and 4 [8]. New work using labor supply data puts the value of r* slightly below 1, and rules out values greater than 2 [9]. Variances are difficult to estimate even with large data sets in health care because of the large coefficients of skewness and kurtosis in medical expenditure data, and the requirement of having data representing expenditure distributions for people with similar or identical insurance coverage and health risks.

with ECEA—would still make it impossible to rank one intervention above another except in cases of pure dominance (one is better than the other on all measures of value), and that becomes increasingly unlikely as more and more elements of value are added to the model. A second approach is to selectively, and systematically, add elements to either the numerator (net cost) or denominator (net benefit as measured in QALYs or utility), as most appropriate. This approach has the advantage of retaining a cost per outcome ratio form, but is likely to affect what threshold is applicable. A third approach is to monetize all benefits so that a net monetary benefit (NMB) can be calculated. However, many health policy analysts and their clients find the step of monetizing health and related benefits objectionable. Indeed, that very objection led to the adoption of CEA over the previously-developed CBA approach to valuing health-related investments. NMB accomplishes the aggregation across different dimensions of value by using the same potentially objectionable step of monetizing each type of benefit. A fourth approach solves this problem by providing a way to aggregate the multiple elements of value into a single, non-monetary, metric: the topic we consider next.

**5.5 Multi-Criteria Decision Analysis (MCDA)**

Neither ECEA nor ACEA provide a single measure of value, since they cannot fully aggregate across the various dimensions of value that they describe. One approach does provide such aggregation: Multi-Criteria Decision Analysis (MCDA), building on the work of Keeney, Saaty, von Winterfeldt and Edwards, Barron and Edwards and others. [10-13]. In general, these approaches elicit from "the decision maker" the tradeoff values to incorporate issues that cannot or have not been included in CEA or CBA. Thus, these models provide a unified single dimensional measure of value of alternative choices using a multi-attribute metric that combines the preference weights specified by "the decision maker" and the performance of alternative "candidates" along each of the dimensions of value. Although other approaches exist, the two most common MCDA methods used in health care are the Analytic Hierarchy Process (AHP) and Multi-Attribute Utility Theory (MAUT). In the simplest form (linear MAUT), the value score is a weighted linear sum of each candidate's performance along each relevant dimension. More complex models use multiplicative models. The largest differences between these approaches appear in the processes used to elicit the decision makers' value structures.

The key components of these MCDA models are the value weights supplied by the decision maker(s) and measures of performance of each intervention candidate along each specified criterion. In effect, MCDA models attempt to re-create a systematic "utility function" of the decision maker that formally expresses tradeoffs between attributes of medical interventions that create value—"goods" in the usual economic sense. However, rather than estimating them from observed behavior, as economists might normally do, they seek to elicit the tradeoffs through various structured processes. These weights will likely differ considerably across different stakeholders (e.g., provider, payers, or patients) and may well differ even within any of these groups.

MCDA models have a series of specific advantages compared to less-formal processes. [14]. In particular, using formal MCDA models:

- Makes the decision making process more transparent;
- Guides investment in data improvement (to where it most affects decisions);
- Allows a "test drive" of alternative program specifications in multi-dimensional value space;
- Allows "reverse engineering" to improve product specifications;
- Bypasses many cognitive errors made in intuitive human judgments; and
- May assist in decision convergence for divergent interests.

To be clear, we believe that specific value measures (such as QALYs) are necessary core elements in any broader MCDA model of value in health care. In some cases, these elements may dominate the model, while in other decision contexts, other attributes of the choices may assume greater importance (larger weights). We view MCDA primarily as a method to expand upon the elements and dimensions of value that are not readily captured in CEAs. The most obvious "omitted element" concerns issues of distribution (otherwise known as "fairness," "equity" or "disparities"). Other important elements may enter these models, particularly when they are difficult or impossible to capture in traditional CEA models. Section 5.2 discussed some of these issues in greater detail.

MCDA models readily allow analysts to employ different perspectives while using the same basic model, merely by changing which dimensions of value (criteria or attributes) are included in the model and the weights they receive in the value structure. Consider five of the potential perspectives outlined in Section 2: the Public (societal), Providers (doctors, hospitals, etc.), Payers (insurers), Patients, and Producers (those who create and produce medical drugs and devices). Each might choose different dimensions of value, or could share some, but with different weights. Others with more defined self-interests (providers, producers, payers) would likely have differing perspectives [14].

Different stakeholders may well have different perspectives (and hence weights) in an MCDA, but the proper use of two perspectives seems clear to us. First, MCDA models should always be presented—at a minimum—using a societal or payer perspective for the same reasons that the societal or payer perspective should always be presented as one option in CEA models. In a specific example, health plans (either public or private) should use a societal or payer perspective in determining coverage of medical interventions (what is covered and at what cost to the patient) and reimbursement to providers. But additional (separate) perspectives can also be developed within the same MCDA structure.

Second, once a health plan determines coverage and reimbursement, a separate viewpoint may be useful—that of the individual patient. Here, MCDA models using individual patient-specific value weights could help patients choose among available medical interventions—those chosen for coverage by their health plan—using both their own value weights and their out-of-pocket costs. Such formal models might help patients choose among complex alternatives such as alternative chemotherapies for cancer, surgical vs. medical intervention for some disorders, or various strategies for mental illnesses. And in creating MCDA-to-cost ratios comparable to cost-effectiveness ratios in CEA, a societal view and the patient view could readily use different measures of cost.

MCDA models differ from ECEA models in two ways. First, ECEA incorporates only a fixed set of criteria beyond health gains and social cost—the dimensions of protection against individual risk and issues of distribution (equity). MCDA in concept could incorporate not only these but other dimensions of value for decision makers. Second, and most distinctly, ECEA (and ACEA as well) does not provide a basis for assessing tradeoffs between gains in these desirable criteria, whereas MCDA explicitly elicits decision-makers' preferences and uses them to provide a single-dimensioned measurement of value. From this perspective, ECEA and ACEA can be seen as a subset of MCDA, but both lacking both flexibility in the number of criteria included in the depiction of value, and a method for measuring tradeoffs between desirable criteria.

Another use of MCDA allows incorporation of things considered valuable to the decision maker but which cannot readily be accommodated in standard CEA, as Section 5.2 discusses. In MCDA, at least

some of these can be incorporated using subjective measures of how well each competing technology performs on these other dimensions of value. For example, in a vaccine program, one could state the "fit" with existing vaccine schedules on (say) a 0-10 scale, and incorporate that "fit" attribute with a chosen weight. *In concept,* one could incorporate this in elaborate measures of the cost of introducing alternative vaccine programs, but in practice, such detail may be impossible.

MCDA can capture other issues that remain beyond the attainable data for a full CEA or ACEA model. Most of these fall into the category of states of the world where we have no existing utility measurements (and where analysts cannot reasonably expect to acquire such measurements). Consider as an example the "fear factor" from a highly dangerous and virulent disease (such as Ebola or Zika). MCDA can incorporate subjective judgments about the ability of technologies to avoid such fear (with an appropriate weight included in the model) in lieu of attempting to measure population-level utilities when people are confronted with varying degrees of risk from such diseases. However, such utilities, if available, could be included in an ACEA model.

Existing MCDA models suffer from a number of curable defects, detailed further in Box 1. Some of these affect only AHP, while others affect only MAUT, the two most commonly used measurement models in healthcare decision support. We view none of these defects as "fatal," and urge continued research to address these issues to improve MCDA performance and ease of use.

***[Box 1 about here]***

**5.6 Budget Setting and Resource Allocation with MCDA.**

MCDA models suffer from a remaining problem. They do not currently provide a clear method for making "investment choices", e.g., deciding which new technologies to add to the health benefit package. And in general, they provide no basis for guiding budget-setting itself, in contrast to CBA and CEA.

Consider first the standard rule of CBA: invest in or include all projects where the benefit-to-cost ratio exceeds 1. In CEA, the standard rule says to invest if the cost per health unit gained (e.g., QALY or DALY) is smaller than the decision-makers' pre-determined cutoff value (and in parallel, reduce investments in interventions with cost-effectiveness ratios exceeding the predetermined cutoff). The World Health Organization (WHO) has recommended using a decision cutoff of 1X to 3X per capita Gross Domestic Product (GDP), although it has more recently been reviewing alternative threshold estimation approaches [17]. The English National Health Service uses a cutoff of £20,000 per QALY, increasing to £30,000, and even to £50,000 in special circumstances. Other alternative approaches lead to lower or higher cutoff values (e.g., see Neumann et al., Section 1.3.3) [18].

CEA measures have and can be used to help resource allocation decisions without formally using a specific cutoff. The common use of "league tables" to help people understand where a medical intervention or technology "fits" in the general realm of common use and practice can inform some decisions about health technology use. In these settings, an implicit CEA cutoff is used in lieu of an explicit measure. This approach, however, tacitly assumes that existing "acceptable" interventions fall within some acceptable range of CEA values.

In the world of MCDA models, one approach to prioritize investments assumes that an exogenously-given investment budget is available, and recommends investing in projects with the highest multi-attribute utility score first, and proceeding down the list of available projects until the investment budget is consumed [19]. But, as noted, this approach provides no guidance as to how large the investment budget should be.

A recent ISPOR Good Practice Task Force report on MCDA addressed the problem of budget-setting (or the parallel problem of how to recommend the best investments), discussing two specific options. These approaches either (a) include "cost" in the MCDA model with a specific (negative) weight, or (b) measure the MCDA value of interventions thought to be susceptible to being removed from use [20]. That Task Force concluded that neither of these approaches was wholly satisfactory, and recommended further research.

To be clear, we recommend that MCDA models do *not* include costs as a (negative) attribute. As noted, using "cost" as an attribute was one of the approaches considered in the earlier ISPOR Task force assessing MCDA. That report concluded (and we agree) that using costs as an MCDA attribute implies knowledge of the appropriate willingness to pay (WTP) for the value bundle, yet the MCDA framework provides no basis for making such judgments. [20]. Instead, users of MCDA models should treat them as a comprehensive benefit (value) measure, independent of cost. Then (as we discuss shortly) research must develop the best methods to determine the investment budget or an acceptable threshold of MCDA value-to-cost ratio to guide investment decisions.

One new approach has subsequently been suggested [16]. This approach builds upon the availability of a decision cutoff in a CEA context (such as $100,000 per QALY) and provides potential guidance for MCDA models that contain QALYs as an important attribute of value. In brief, this approach scales the CEA "cutoff value" upward to account for other dimensions of value not measured in the CEA framework. Thus, for example, if QALYs accounted for two-thirds of the weight (or value) in an MCDA model, and the CEA cutoff value was $100,000 per QALY, then the MCDA value-to-cost cutoff would be 3/2 times the CEA cutoff, or (in this example) $150,000 per unit of MCDA value. In effect, agreed-upon values for a CEA cutoff (such as may exist) serve as a *numeraire* to estimate the total value of the expanded multi-criteria value index. By analogy, it would be like adding new safety features to automobiles and then re-evaluating them taking the new features into account. The overall WTP for any automobile should expand with new safety features added, but the rankings of value would shift depending on how much new safety-enhancement was built into any particular model, and each vehicle would then be ranked on the ratio of its new MCDA value to its cost.

In some settings, a separate issue arises: non-alignment between budgets allocated for provision of health services and the decision rule to determine acceptability of medical interventions. First, consider this issue in the familiar context of CEA analysis. It might arise, for example, when a prototypical 'Minister of Health" establishes the CEA cutoff, while a "Minister of Finance" establishes the budget for the health care system. Even if initially in balance, new technologies or health care interventions might emerge that had an "acceptable" CEA cutoff, yet their inclusion into the package of services offered to enrollees might greatly exceed the budget. Drugs to treat hepatitis C appear to have created this issue in some situations, and (for example) the emergence of an effective treatment for Alzheimer's Disease could well cause it to appear in the future. This issue has been described as the problem of "affordability". In this situation, the budget creates an implicit cost-effectiveness cutoff that is more stringent than the official cutoff. Economists would call this the "shadow price."

In the long run, the most desirable solution would bring the budget setting process and the CEA-cutoff determination (or an expanded MCDA-cutoff determination) into alignment. The official cutoff and the shadow price should be one and the same. Achieving that alignment, however, may be difficult to impossible, at least in the short run. When a misalignment occurs, it would appear that any of four logical solutions are available (none exclusionary to the others): (1) increase the budget;[2] (2) tighten the official CEA (or MCDA) cutoff, (3) cut back on or eliminate interventions with excessive CEA ratios (or in an MCDA world, with poor MCDA value-to-cost ratios) or (4) as a variant on (2), estimate the shadow cutoff value implied by the budget, and use that as the "real" CEA (or MCDA) cutoff. In situations such as these, resolving the lack of alignment between cutoff rules and budgets must—at least in the short and intermediate run—account for costs of change and other issues discussed previously in Section 4 [3].

This issue could become more prominent with the adoption of MCDA value measures, particularly when they expand measures of value beyond traditional health outcomes. This could place more pressure on budgets than would arise in a pure CEA-driven world. If the cutoff value should expand in MCDA models beyond that used for CEA, then (other things equal) budgets will commonly have to expand, to accommodate the costs of acquiring the additional elements of value. Some of the added value (as measured by the MCDA approach) may well lie outside the domain of health, e.g., elements of value arising in the domains of education, social justice, or others. Budgets of health care providers or payers would not normally incorporate these "other" issues, hence exacerbating non-alignment between budgets and pre-specified MCDA-cutoffs. Offsetting these potential increases, some existing interventions might fall in overall value as measured by the MCDA approach than previously ranked using only CEA. In this situation, pruning out some interventions that score poorly on the MCDA value-to-cost index could relieve budgets.

None of these solutions are easy to achieve in real-world situations. In many cases, CEA ratios are not available for the entire portfolio of existing interventions, so deciding where to "cut back" cannot rely entirely on using incremental CEA ratios. This would happen even more frequently where an MCDA model was employed, if for no other reason than the more extensive data demands of MCDA (beyond those of CEA) would make it less likely that all interventions in the portfolio had been evaluated using the appropriate metric. In situations such as these, the most likely outcome is to turn to deliberative bodies to resolve the dilemma, a topic to which we turn to next.

Another approach exists once an MCDA metric is chosen: allow people to vote on the appropriate cutoff for the bundle of benefits as defined by a particular MCDA weighting structure. In this setting, people would "vote" with a number for the cutoff value to define an acceptable investment. For a number of reasons, the median value is the most appropriate to use in such a vote [21,22]. The same approach, of course, could be used directly to choose a CEA cutoff through population voting.

### 5.7 Deliberative Processes

[2] Future budget expansion may not be inevitable. One could imagine cases where a high expenditure for an intervention "today" reduced subsequent-year budgets, e.g., through elimination of a contagious disease, or for reduced future liver transplant costs from lower alcoholic cirrhosis. Proper intertemporal budgeting with discounting would accommodate such issues, but politically set budgets commonly do not have the appropriate temporal horizon, and thus cannot accommodate such issues readily.

A deliberative process is characterized by the careful, deliberate consideration and discussion of the advantages and disadvantages of various options [23]. Hence, a deliberative process is best considered as an aid to thought and judgment. Properly executed, it will be—compared with an *ad hoc* process—more comprehensive in the relevant issues embraced, more consistent in the way they are embraced, and more engaging of the people affected by the outcome. In making health care resource allocation decisions, deliberative processes have been useful in incorporating wider dimensions of value, such as: (a) social and cultural values, such as other types of outcome beyond health gain; (b) other social and personal values not typically taken into account, such as issues of equity and fairness; and (c) practical issues of operational feasibility [24].

Currently, we have identified several examples of deliberative processes in regular use. For example, the method used by the National Institute for Health and Care Excellence (NICE) for evaluating healthcare technologies in England and Wales includes formal submissions from interested parties, as well as consultations and invited commentaries from consultees and commentators, systematic reviews, technical modeling exercises, and multiparty representation in the (large) deliberative committee that hears witnesses [25]. Rawlins and Culyer and Rawlins et al. discuss how this process has been used to incorporate several social judgments, including equity considerations, into NICE's decisions [26,27].

Also, the Scottish Medicines Commission uses several "modifiers" in appraising new medicines, which are sometimes used to justify accepting a higher cost per QALY (see Section 4) [3]. In Australia, George et al. argue that the Pharmaceutical Benefits Advisory Committee (PBAC) has taken account of factors other than cost-effectiveness in reaching its reimbursement decisions, including severity of disease, the absence of other effective therapy, and the financial burden on the patient if the therapy was not reimbursed [28]. Examples of bodies in the US that use deliberative processes to incorporate other considerations when formulating policy recommendations are the Advisory Committee on Immunization Practices and the U.S. Preventive Services Task Force [29,30].

Although the evidence is not strong, the possible advantages of deliberative processes are: (i) decision-makers acquire a better grasp of the strengths and weaknesses of the underlying cases and can better defend their decisions; (ii) consensus building is enhanced; (iii) the revelation of evidence gaps helps to inform downstream research programs; (iv) stakeholders and their peers are more likely to accept and implement decisions that they have had a hand in shaping; (v) possible selection bias through the membership of decision-making panels becomes relatively more obvious; and (vi) context-free evidence can be re-interpreted in relevant contexts [24]. Daniels and Sabin argued that the reasonableness or legitimacy of the process was crucial to obtaining acceptance of certain health plan decisions [31]. It can be argued that a transparent deliberative process can increase the legitimacy of decision making compared to a "black box" process where it is unclear what has been considered and how the decision was reached.

On the other hand, deliberative processes may be relatively informal and unstructured. Thus, without a formal set of prompts, key issues may be overlooked, decisions may be reached in an unstandardized way, and the biases of decision makers may not be adequately contested. There may also be a lack of quantification of elements other than health system costs and health gain, which might in principle be taken into account. For these reasons, it may be valuable to deliberative bodies to incorporate an explicit framework such as MCDA. Some HTA bodies or analysts, such as Institute for Clinical and Economic Review (ICER) in the U.S., have used this approach. As in all such uses, MCDA analyses do not *make* decisions. Rather, they inform the process and help quantify thinking.

If a deliberative process were to use MCDA to help structure its decisions, then that deliberative body must set the weights for the MCDA model in advance. Otherwise, the MCDA model has no meaning in the context of the deliberation. Whether used to augment deliberative processes or as stand-alone processes, using MCDA models in the context of group decision making requires voting or ranking procedures for groups to choose among alternatives. Box 2 contains a brief discussion of some pertinent issues for these "choice-related" procedures.

***{Box 2 about here: voting and ranking methods]***

Deliberative processes, no matter what voting methods are used, have participants integrating a considerable amount of possibly complex information to reach a single conclusion. As is now increasingly understood in the field of behavioral economics, human decision making is often distorted by how questions are framed, inaccurate assessment of probabilities, and other important human fallibilities [36]. The formal structure of MCDA may avoid at least some of these issues that may persist in less-structured deliberative processes.

**5.8 Other Value Elicitation Methods.**

The previous section discusses use of MCDA models to aggregate multiple dimensions of value into a single metric. The commonly used models such as MAUT and AHP have specific methods embodied in them to elicit the needed value tradeoffs—primarily arising from the work of people in the field of decision science. These methods have their own strengths and weaknesses (see Box 1), but other approaches may prove fruitful, some of which arise from the methods of economics.

Because health itself cannot be bought and sold in a market setting, and because health care markets are distorted by insurance or government subsidies, direct valuation of a health intervention or a health insurance product as a differentiated good through observed market prices—as economists might normally do—is difficult. In a recent paper, Basu and Sullivan discussed the rationale of using stated preference methods for developing "hedonic" value frameworks for health insurance products to inform the decision on whether a product should be covered or subsidized by insurance, given its price [37]. They propose that discrete choice experiments (DCE) in a nationally representative sample be used to elicit WTP for health insurance products that would cover a new health intervention with specific attributes. These elicitations should be carried out among both the patients who are the direct beneficiaries of this intervention as they are diagnosed with the specific clinical condition and also healthy individuals who do not have that specific clinical condition, but face varying risks of being diagnosed with that condition in a given year. This combination of values from both the patients and non-patients would reflect the true value of a healthcare intervention as it would incorporate the value of health insurance covering that intervention [38]. Based on these elicitations, a WTP value index could be developed relating to each of the dimensions of a health care intervention.

This value index will not reflect marginal value at market equilibrium, as would appear in hedonic pricing estimates, but it would capture the distribution of marginal value in the population. This approach can inform two specific decision-making processes: (1) whether a health insurance budget could be expanded to cover a new technology, and (2) which existing technology can be displaced to accommodate the new technology that maximizes value in the health plan population. If either or both cannot accommodate the budget impact of the new technology, then coverage may not be feasible.

Basu and Sullivan lay out a research agenda that can help develop this concept of hedonic valuation for health care value frameworks [37].

**5.9 Summary and Recommendations**

No existing method completely or perfectly solves the two aggregation problems that we consider in this section. It is widely understood that CEA, CBA, and related models cannot deal with aggregation across individuals (equity, fairness, disparities, etc.), and in actual practice (if not in concept), cannot incorporate some important components of value (and occasionally components of cost). Consensus-building approaches (i.e., "deliberative methods") are often unstructured, idiosyncratic, and easy for individuals to influence or control through agenda manipulation, power of status or personality. ECEA and even further expansions of such models (such as ACEA) gather information about additional outcomes beyond those considered in standard CEA, but provide no mechanism for aggregating across the relevant elements or dimensions of value. They provide uncontestable answers to questions about coverage of an intervention only when pure dominance emerges: i.e., one candidate exceeds another on every possible dimension of value. Thus, these are not satisfactory methods to carry into the future without further effort to improve upon them.

Public and private insurers use deliberative decision making for payer coverage and reimbursement decisions. A transparent deliberative process can increase the legitimacy of decision making. Currently, such processes often lack transparency, and it is unclear what factors have been considered and how decisions were reached. Deliberative processes today are often informal and unstructured. Thus, without a standardized approach, key issues may be overlooked, decisions may be reached in an unstandardized way, and potential biases of decision makers may not be adequately explored. This can be improved, both by the introduction of cost-per-QALY evidence and by the use of more structured decision making to take account of preferences about the weight to be given to health gain—e.g., about disease severity, equity of access, or unmet need. A systematic comparison of the processes used by various private and public pharmacy and therapeutic committees and health technology assessment programs may further help others determine best practices for their own setting.

MCDA models may provide the best opportunity for improvement, but they have not yet been perfected. To improve these methods, we urge progress on two fronts. First, we must expand the use of MCDA models in real-life decision settings and learn from these experiences how well they work. We may learn that they seldom differ from standard CEA in the investment advice they give, or we may learn that formally incorporating these "other issues" importantly changes many decisions. We cannot know until we "run the experiment."

Second, we also need more research on key aspects of MCDA modeling and use. Just as the current "gold standard" of CEA did not begin in its current form, we can expect that MCDA will evolve for the better in the future. In CEA, both for lack of the proper conceptual framework and the lack of data, earlier efforts focused on things that were easily measurable. These first included "deaths averted" or similar measures, and soon expanded to "life years saved" or similar measures. Then came the notion of quality adjustment, leading to the current metric of the QALY based on work that relies on population-based estimates of reported quality adjustments for different health-related conditions. A similar but not identical measure—DALYs—is used by WHO, the World Bank, and others: for this measure, expert judgment has been used (at least initially) rather than population survey data to create the adjustment factors [17], although DALY estimates increasingly turn to population-based metrics when available. A third related measure is the capabilities-adjusted life year, using a capabilities index

instead of a health state utility, initially proposed by Amartya Sen, progressed by Nussbaum, and operationalized by Anand et al. [39,40].

We likely stand at a similar point in the evolution of MCDA models with important issues to resolve before they reach their full potential. Some have issues in ease of use. Some have methodological flaws such as the risk of "rank reversal" as new technology options emerge. Little is known about important human factor issues associated with the use of various MCDA models, including ease of use, susceptibility to strategic manipulation, and ease of comprehension of the methods (and hence acceptance of the results). And finally, these approaches are quite data-intensive compared with CEA modeling since they require measuring each candidate technology on multiple dimensions of value rather than the single dimension of QALYs.

We recommend greater testing and use of MCDA models, pushing the frontiers of their use and continuously comparing their results with those of standard CEA and similar models. Using ECEA or even more broadly—ACEA—models may provide a halfway-house step to MCDA by facilitating data acquisition and refinement. But ECEA provides no way to combine multiple dimensions of value into a single index of merit. MCDA provides the logical basis for this next step. An important missing element is a universally acceptable method to elicit value weights. Current approaches (see Box 1) have known defects, but have nevertheless been demonstrated to assist in decision making in complex health care settings but further improvement is needed to bring MCDA models to full flower.

What are the best methods for acquiring value weights, particularly in settings with groups acting as decision makers? AHP and MAUT offer different approaches with different strengths and weaknesses. DCEs using representative populations offer another approach to establishing proper weights [37]. Other approaches may emerge as well. All should be tested and compared both for methodological soundness and human factors (ease of use, etc.).

Acknowledgements: The authors did not receive any funding for this work other than reimbursement from ISPOR for travel expenses, as needed, for two Special Task Force meetings. Darius Lakdawalla is a consultant to Precision Health Economics (PHE) and an investor in its parent company, Precision Medicine Group. PHE conducts research for firms in the life sciences industries.

Ethics Statement: The authors state that no ethical approval was needed.

Disclaimer: The views expressed in this article are the personal views of the authors and may not be understood or quoted as being made on behalf of or reflecting the position of the agencies or organizations with which the authors are affiliated.

**References**


1. Bator, FM. The simple analytics of welfare maximization. American Economic Review 1957; 47(1):22-59.
2. Garrison LP, Pauly MV, Willke RJ, Neumann PJ. An Overview of Value, Perspective, and Decision Context—A Health Economics Approach: An ISPOR Special Task Force Report Value Health. 2018;21(2):XXXX
3. Danzon P, Drummond M, Towse A, Pauly MV. Objectives, Budgets, Thresholds, and Opportunity Costs—A Health Economics Approach: An ISPOR Special Task Force Report Value Health. 2018;21(2):XXXX.
4. Lakdawalla D, Doshi JA, Garrison LP, et al. Defining Elements of Value in Health Care—A Health Economics Approach: An ISPOR Special Task Force Report. Value Health. 2018;21(2):XXXX.
5. Verguet S, Kim JJ, Jamison DT. Extended cost-effectiveness analysis for health policy assessment: a tutorial. Pharmacoeconomics 2016;34(9):913-23.
6. Arrow, K. J. (1965). "Aspects of the Theory of Risk Bearing in The Theory of Risk Aversion." Helsinki: Yrjo Jahnssonin Saatio. Reprinted in: Essays in the Theory of Risk Bearing, Markham Publ. Co., Chicago, 1971, 90–109.
7. Pratt JW. Risk aversion in the large and in the small. Econometrica, January-April, 1964, 32(1-2): 122-136.
8. Garber AM, Phelps CE. Economic foundations of cost-effectiveness analysis. Journal of health economics. 1997 Feb 28;16(1):1-31.
9. Chetty, R. A new method of estimating risk aversion. American Economic Review 2006; 96(5):1821-1834.
10. Keeney RL, Raiffa H. Decisions with Multiple Objectives: Preferences and Value Tradeoffs. New York: Wiley, 1976: Reprinted in 1993 by Cambridge University Press.
11. Saaty TL. The Analytical Hierarchy Process, Planning, Priority. Resource Allocation. RWS Publications, USA, 1980.
12. von Winterfeldt D and Edwards W (1986). Decision Analysis and Behavioural Research. Cambridge University Press: New York.
13. Edwards W, Barron FH. SMARTS and SMARTER: Improved simple methods for multiattribute utility measurement. Organizational Behavior and Human Decision Processes 1994; 60:206-325.
14. Phelps CE, Madhavan G. Using multi-criteria approaches to assess the value of health care. Value Health. 2017;20(2):251-5.
15. Barron FH, Barrett BE. Decision quality using ranked attribute weights. Management Science 1996; 42(11):1515-1523.
16. Phelps CE, Madhavan G. Resource allocation in decision support frameworks. Cost Effectiveness and Resource Allocation 2017; 15(S1), in press.
17. Tan-Torres Edejer T, Baltussen RM, Adam T, Hutubessy R, Acharya A, Evans DB, Murray CJ. Making choices in health: WHO guide to cost-effectiveness analysis, 2003.
18. Neumann PJ, Sanders GD, Russell LB, Siegel JE, Ganiats TG. Cost-effectiveness in Health and Medicine. New York: Oxford University Press, 2016.
19. Phillips LD, Bana e Costa CA. Transparent prioritisation, budgeting and resource allocation with multi-criteria decision analysis and decision conferencing. Annals of Operations Research 2007; 154:(1): 51-68.
20. 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 in Health 2016; 19:125-137.

21. Galton F. Vox populi. Nature, March 7,1907; 450-51.
22. Balinski M, Laraki R. Majority Judgment: Measuring, Ranking, and Electing, Cambridge, MA: MIT Press, 2010.
23. Hajer MA, Wagenaar H (eds) Deliberative policy analysis: Understanding governance in the network society. Cambridge: Cambridge University Press, 2003.
24. Culyer AJ, Lomas J. Deliberative processes and evidence-informed decision making in health care: do they work and how might we know? Evidence and Policy: A Journal of Research, Debate and Practice 2006; 2(3): 357-371.
25. National Institute for Health and Care Excellence (NICE). A guide to the methods of technology appraisal. London, July 2003. Available at: http://www.nice.org.uk/aboutnice/howwework/devnicetech/guidetothemethodsoftechnologyappraisal.jsp (accessed 5 December 2013).
26. Rawlins MD, Culyer AJ. National Institute for Clinical Excellence and its value judgments. British Medical Journal 2004; 329: 224-227.
27. Rawlins, M et al. Pharmacoeconomics: NICE's approach to decision making. British Journal of Clinical Pharmacology 2010; 70: 346-349.
28. George B, Harris A, Mitchel A. Cost-effectiveness analysis and the consistency of decision-making: evidence from pharmaceutical reimbursement in Australia (1991-1996). Pharmacoeconomics 2001; 19: 1103-1109.
29. Smith JC. The structure, role and procedures of the US Advisory Committee on Immunization Practices (ACIP). Vaccine 2010; 285: A68-A75.
30. Agency for Healthcare Research and Quality. US Preventive Services Task Force Procedure Manual. AHRQ Publication No. 08-05118-EF.
31. Daniels N, Sabin, JE. Accountability for reasonable priority setting. BMJ 2008; 9,337,a1850.
32. Arrow KJ. A difficulty in the concept of social welfare. Journal of Political Economy 1950; 58(4):328-346.
33. Gibbard A. Manipulation of voting schemes: A general result. Econometrica 1973, 41 (4): 587–601.
34. Satterthwaite MA. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. Journal of Economic Theory 1975; 10 (2): 187–217.
35. Plott CR, Levine ME. A model of agenda influence on committee decisions. American Economic Review 1978; 68(1):146-160.
36. Kahneman D. Thinking, Fast and Slow. Macmillan. 2011 Oct. 25.
37. Basu A, Sullivan SD. Towards a hedonic value framework in health care. Value Health 2017; 20(2):261-5.
38. Lakdawalla D, Malani A, Reif J. The insurance value of medical innovation. J Public Economics 2017; 145: 94-102.
39. Nussbaum M. Capabilities as fundamental entitlements. Feminist Economics 2003; 9: 33-59.
40. Anand P, Hunter G, Carter I, Dowding K, Guala F, Van Hees,M. The development of capability indicators. Journal of Human Development and Capabilities 2009; 10: 125-152.