

Chapter 2

Theoretical Foundations of MCDA

Dean A. Regier and Stuart Peacock

Abstract Decision-makers in the healthcare sector face a global challenge of developing robust, evidence-based methods for making decisions about whether to fund, cover, or reimburse medical technologies. Allocating scarce resources across technologies is difficult because a range of criteria are relevant to a healthcare decision, including the effectiveness, cost-effectiveness, and budget impact of the technology; the incidence, prevalence, and severity of the disease; the affected population group; the availability of alternative technologies; and the quality of the available evidence. When comparing healthcare technologies, decision-makers often need to make trade-offs between these criteria. Multi-criteria decision analysis (MCDA) is a tool that helps decision-makers summarize complex value trade-offs in a way that is consistent and transparent. It is comprised of a set of techniques that bring about an ordering of alternative decisions from most to least preferred, where each technology is ranked based on the extent to which it creates value through achieving a set of policy objectives. The purpose of this chapter was to provide a brief overview of the theoretical foundations of MCDA. We reviewed theories related to problem structuring and model building. We found problem structuring aimed to qualitatively determine policy objectives and the relevant criteria of value that affect decision-making. Model building theories sought to construct consistent representations of decision-makers' preferences and value trade-offs through value measurement models (multi-attribute value theory, multi-attribute utility theory, and the analytical hierarchy process), outranking (ELECTRE), and reference (weighted

D.A. Regier (✉)

Canadian Centre for Applied Research in Cancer Control (ARCC), Cancer Control Research, British Columbia Cancer Agency, Vancouver, BC, Canada

School of Population and Public Health, University of British Columbia, Vancouver, BC, Canada

e-mail: dregier@bccrc.ca

S. Peacock

Canadian Centre for Applied Research in Cancer Control (ARCC), Cancer Control Research, British Columbia Cancer Agency, Vancouver, BC, Canada

Faculty of Health Sciences, Simon Fraser University, Burnaby, BC, Canada

and lexicographic goal programming) models. We conclude that MCDA theory has largely been developed in other fields, and there is a need to develop MCDA theory that is adapted to the healthcare context.

2.1 Introduction

Decision-makers in the healthcare sector face a global challenge of employing robust, evidence-based methods when making decisions about whether to fund, cover, or reimburse medical technologies. Historically, technology assessment agencies promoted cost-effectiveness analysis as the primary decision aid for appraising competing claims on limited healthcare budgets. The recommended metric to summarize cost-effectiveness, the incremental cost-effectiveness ratio (ICER), has the incremental costs of competing technologies in the numerator and quality-adjusted life years (QALY) gained in the denominator (Drummond 2005). Judgment surrounding value for money is determined against a cost-effectiveness threshold, which represents the opportunity cost to the healthcare sector of choosing one technology over another. While cost-effectiveness analysis is a necessary component when informing resource-constrained decisions, it is not a sufficient condition. This is because a range of criteria are relevant, including the incidence, prevalence, and severity of the disease; the population group affected; the availability of alternative technologies; the quality of the available evidence; and whether the technology contributes technological innovation (Devlin and Sussex 2011). Technology assessment agencies have gone as far as publishing the criteria they consider in their decision-making contexts, including how each criterion may be reflected in funding decisions (NICE 2008; Rawlins et al. 2010).

Publishing the criteria that decision-makers consider when setting priorities can provide clarity for stakeholders on how committees make decisions (Devlin and Sussex 2011). Making explicit the extent to which decision criteria influence program funding will enhance the legitimacy, transparency, and accountability of decisions and will encourage public trust in the decision-making process (Regier et al. 2014a; Rowe and Frewer 2000). Further, it will improve the consistency of decisions, can provide an opportunity for decision-makers to engage the public, and can serve to sharpen the signal to industry about what aspects of innovation are important and where research and development efforts should be directed (Devlin and Sussex 2011).

Multi-criteria decision analysis (MCDA) represents a set of methods that decision-makers can use when considering multiple criteria in priority-setting activities. It is a decision aid that helps stakeholders summarize complex value trade-offs in a way that is consistent and transparent, thus leading to fairer decision-making (Peacock et al. 2009). MCDA makes explicit the criteria applied and the relative importance of the criteria. As such, MCDA is a process that integrates objective measurement with value judgment while also attempting to manage subjectivity

(Belton and Stewart 2002). To enable these goals, the theories and methodologies supporting MCDA need to allow for both the technical and nontechnical aspects of decision-making; they need to include sophisticated quantitative algorithms while also providing structure to a decision-making process. All help to promote the replicability and transparency of policy decisions (Belton and Stewart 2002).

The purpose of this chapter is to provide an overview of the theoretical foundations of MCDA. Particular attention is given to two key parts of MCDA: problem structuring and model building. Problem structuring refers to determining policy objectives through methods that illuminate the policy-relevant criteria. Model building requires constructing consistent representations of decision-makers' values and value trade-offs. MCDA theory in healthcare is under-researched - for this reason, we draw on theories developed in other disciplines in the following sections.

2.2 Principles of MCDA and Decision-Making

MCDA begins with decision-makers encountering a choice between at least two alternative courses of action. A key principle of MCDA is that decision-makers consider several objectives when judging the desirability of a particular course of action (Keeney and Raiffa 1993). It is unlikely that any one program will satisfy all objectives or that one course of action dominates another. Each program will meet the policy objectives at different levels, and trade-offs will be inherent when making decisions (Belton and Stewart 2002). Policy objectives can be thought of as the criteria against which each program is judged. Choosing one program over another will entail opportunity cost. The decision is one where multiple objectives (criteria) need to be balanced and acceptable. The aim of decision-makers is to make the best choice between alternative courses of action that are characterized by multiple criteria (Belton and Stewart 2002). Doing so can help ensure that decisions are consistent with the policy objectives.

These key principles suggest several basic assumptions of MCDA. First, decisions are made under constrained resources – not all programs can be funded and choosing one program over another will entail opportunity cost. Second, decision-makers' objectives are within their personal discretion and are not normatively determined by theories from ethics or economics, such as utilitarianism or social justice (Peacock et al. 2009). Third, a program cannot be thought of as a homogenous “good.” Instead, multiple levels of criteria can describe each alternative program and decision-makers weigh and value each criterion level (Lancaster 1966). Decision-makers can relate the criteria levels to the program alternatives, and incremental changes in criteria can cause a switch from one good with a specific bundle of characteristics to another good with a different combination that is more beneficial. A fourth assumption supporting MCDA value theories is that trade-offs and the relative importance of criteria can be established or that such scores can allow for a rank ordering of programs (Baltussen and Niessen 2006; Baltussen et al. 2006).

2.3 Problem Structuring

Problem structuring is the process of stakeholders identifying policy objectives and decision criteria that they determine are of value (Belton and Stewart 2002). Model building and the use of quantitative methods to determine value were the focus of early MCDA applications. The literature has increasingly acknowledged the importance of problem structuring (Phillips 1984; Schoner et al. 1999). This is due to the recognition that failing to adequately frame and structure the policy problem increases the likelihood of committing a type III error, that is, getting the right answer to the wrong question (Kimball 1957). The theory behind problem structuring begins with understanding the nature of MCDA. Definitionally, MCDA is an aid to decision-making that relates alternative courses of action to conflicting multiple criteria requiring value trade-offs. The decision criteria are determined in relation to decision-makers' objectives. It follows that decision-makers can have differing values with varying sets of objectives and preferences. Decision-makers can dispute which objectives are "right" when choosing between healthcare programs, whether it be to maximize health status, to solve a political problem, or to balance trade-offs between health status and equity. Equally, the solution to any objective is debateable because decision-makers' weightings of criterion will not be homogeneous, possibly leading to different solutions to the problem. Using Ackoff's (1979a, b) lexicon of defining types of decision problems, Belton and Stewart (2010) argue that the MCDA problem can be termed "messy" because both the definition and solution to the problem are arguable. Contributing to the messy categorization is that MCDA criteria can be based on evidence from the hard or soft sciences (objectives are quantitative versus qualitatively assessed) (Goetghebeur et al. 2008). Hester and Adams (2014) defined messy problems as the intersection between hard and soft sciences (Fig. 2.1).

A key component to addressing messy-type problems is the use of facilitation to identify values and frame the multi-criteria problem (Keeney and McDaniel 1992). Following Keeney (1992), decision-makers' core values determine strategic objectives, criteria, and decisions. While decision-makers are likely to know their latent values, their values can change as new information becomes available (Schoner et al. 1999). The goal of facilitation is to translate latent values to make statements regarding the objectives and the set of criteria, the set of alternatives from which to make decisions, and the methods that will be used characterize criteria weights (Belton and Stewart 2002).

Diverging perspectives between decision-makers coupled with system-wide implications suggest that decision-makers may elect to include input from multiple stakeholders. Stakeholders can be identified by focusing on the nature of the health system (Checkland 1981). In soft systems methodology, one framework proposed the following checklist to understand the system and stakeholders under the acronym CATWOE: *C*ustomers are individuals who are directly affected by the system; *A*ctors are individuals carrying out the system activities; *T*ransformation is the purpose of the system; *W*orld View includes the societal purposes of the system; *O*wners are

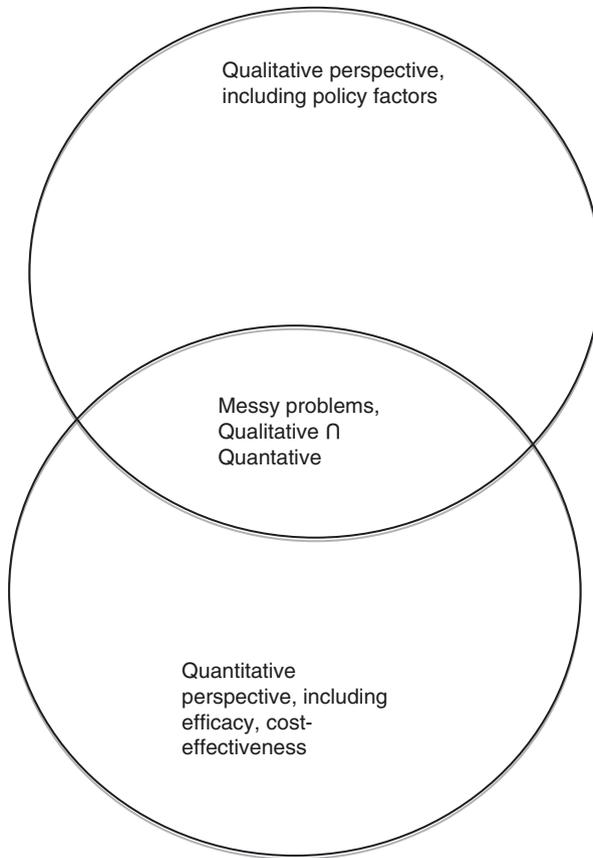


Fig. 2.1 An illustration of messes

those who control system activities; and *Environment* includes the demands and constraints external to the system. Stakeholders can include representatives from government, key decision-makers at institutions, clinicians, healthcare professionals, patients, the lay public, or drug developers. It is emphasized that the inclusion of stakeholders and the extent that stakeholders' views are included in the facilitation process is under the discretion of decision-makers (Belton and Stewart 2002).

Consideration should be given to the idea that stakeholders can exert varying degrees of power to over- or under-influence a particular decision, or to be included as a token participant with little input into the discourse (Arnstein 1969). To mitigate the potential for power structures, facilitation can focus on using deliberative theories that adhere to a process of respectful engagement, where stakeholders' positions are justified and challenged by others, and conclusions represent group efforts to find common ground (O'Doherty et al. 2012; MacLean and Burgess 2010). The facilitator can strive to understand what is going on in the group and

should attend to relationships between participants while being aware that they need to intervene to forward the work of the group (Phillips and Phillips 1993).

Whether or not stakeholders with different backgrounds are included in problem structuring, varying frames will emerge during facilitation (Roy 1996). A frame is a cognitive bias that exists when individuals react differently to a criterion depending on how the information is represented (e.g., number of years gained versus number of years lost) (Belton and Stewart 2010). Through facilitation, decision-makers need to acknowledge frames such that stakeholders can similarly understand the criterion. To do this, Belton and Stewart (2002) identified the following set of general properties to consider, which include domains related to value relevance, measurability, nonredundancy, judgmental independence, completeness, and operationality.

- *Relevance*: Can decision-makers link conceptual objectives to criteria, which will frame their preferences? For example, a criterion may include the cost-effectiveness of a competing program, and an associated objective would be to pursue programs that provide value for money.
- *Measurability*: MCDA implies a degree of measuring the desirability of an alternative against criteria. Consideration should be given regarding the ability to measure or characterize the value of the specified criteria in a consistent manner by allowing for criteria to be decomposed to a number of criteria attribute levels.
- *Nonredundancy*: The criteria should be mutually exclusive with a view to avoid double counting and to allow for parsimony. When eliciting objectives and defining criteria, decision-makers may identify the same concept but under different headings. If both are included, there is a possibility the criteria will be attributed greater importance by virtue of overlap. If persistent disagreement regarding double counting exists in the facilitation process, double counting can be avoided through differentiating between process objectives (how an end is achieved) and fundamental objectives, ensuring that only the latter are incorporated (Keeney and Mcdaniels 1992).
- *Judgmental independence*: This refers to preferences and trade-offs between criteria being independent of one another. This category should be taken in light of preference value functions.
- *Completeness and operationality*: Refers to all important aspects of the problem are being captured in a way that is exhaustive but parsimonious. This is balanced against operationality, which aims to model the decision in a way that does not place excessive demands on decision-makers.

There is a broad literature based on psychology and behavioral economics outlining how judgment and decision-making depart from “rational” normative assumptions (Kahneman 2003). For a comprehensive review of techniques directed at improving judgment, interested readers are referred to Montibeller and von Winterfeldt (2015).

2.4 Model Building

Model building refers to constructing a behavioral model that can quantitatively represent decision-makers' preferences or value judgments. MCDA models originate from different theoretical traditions, but most have in common two components: (1) preferences are first expressed for each individual criterion; and (2) an aggregation model allows for comparison between criteria with a view to combine preference estimates across criteria (Belton and Stewart 2002). The aggregation model establishes a preference ordering across program alternatives. A bird's-eye view of the decision objective(s) is represented through a hierarchical value tree, where there are m criteria at the lowest level of the hierarchy (Fig. 2.2). The broadest objective is at the top. As decision-makers move down the hierarchy, more specific criteria are defined. This continues until the lowest-level criteria are defined in such a way that a clear ordering of value can be determined for each alternative (Belton and Stewart 2002). Generally, aggregation can be applied across the tree or in terms of the parent criteria located at higher levels of the hierarchy.

The following section highlights several underlying theories used in MCDA that support model building through value measurement models, including multi-attribute value theory, multi-attribute utility theory, and the analytical hierarchy process; outranking, focusing on ELECTRE; and reference models using weighted goal and lexicographic goal programming. While the theories described are not comprehensive, we have chosen methods that represent the major theoretical approaches supporting MCDA.

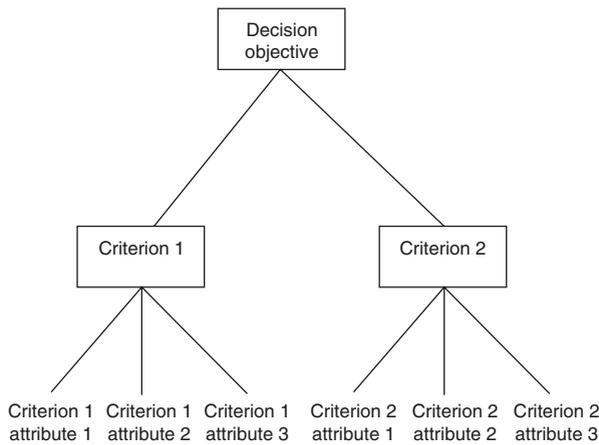


Fig. 2.2 Hierarchical value tree

2.4.1 Value Measurement

Value measurement models aggregate preferences across criterion to allow decision-makers to characterize the degree to which one alternative program is preferred to another (Belton and Stewart 2002). The challenge is to associate decision-makers' preferences to a quantitative real number such that each program alternative can be compared in a consistent and meaningful way. Value measurement approaches have been cited as the most applied methods in MCDA in the healthcare context (Dolan 2010).

2.4.1.1 Multi-attribute Value Theory

The principle that decision-makers want to make consistent decisions is the key building block for multi-attribute value theory. The notion of consistency is formalized through two preference-based assumptions: (i) completeness – any two alternatives can be compared such that alternative a is preferred to alternative b , or b is preferred to c , or they are equally preferred – and (ii) transitivity, if a is preferred to b and b is preferred to c , then a is preferred to c . These assumptions – called axioms in economic utility theory (Mas-Colell et al. 1995) – provide the necessary building blocks for a mathematical proof of the existence of a real-valued function that can represent preferences. That is, these assumptions allow for statements regarding the ability of quantitative values for a program alternative, denoted as $V(a)$ for alternative program a , to represent preferences such that $V(a) \geq V(b)$ if and only if $a \succeq b$, where \succeq is a binary preference relation meaning “at least as good as”; or $V(a) = V(b)$ if and only if $a \sim b$, where \sim denotes indifference between the value of a good. Depending on the context and complexity of the decision problem, the assumptions of completeness and transitivity can be violated in real-world contexts (Camerer 1995; Rabin 1998). In value measurement theory, these axioms provide a guide for coherent decisions but are not applied literally (Belton and Stewart 2002). That is, the axioms are not dogma.

The next set of definitions outline the value score of the criteria. A partial value function, denoted as $v_i(a)$ for program a , is needed such that $v_i(a) > v_i(b)$ when criteria i for program a are preferred to b after consideration of the opportunity costs (also called “trade-offs,” i.e., a sacrifice of one aspect of a good that must be made to achieve the benefit of another aspect of the good). When a performance level of criterion i is defined as attribute $z_i(a)$ for alternative a (i.e., $z_1(a), z_2(a) \dots z_m(a)$) and if the value of a criterion is independent of the other z_i criteria and is increasing in preference, it is denoted as $v_i(z_i)$ (Keeney and Raiffa 1993). Of note, the value of a given configuration that is consistent with these properties is equivalent to $v_i(a) = v_i(z_i(a))$; as such, $z_i(a)$ is called a partial or marginal preference function (Belton and Stewart 2002).

A widely applied aggregation of decision-makers' preferences is the additive or weighted sum approach (Belton and Stewart 2002):

$$V(a) = \sum_{i=1}^m w_i v_i(a) \quad (2.1)$$

where $V(a)$ is the overall value of alternative a , w_i is the relative importance of criterion i , and $v_i(a)$ is the score of a program alternative on the i^{th} criterion (Belton and Stewart 2002; Thokala and Duenas 2012). The partial value functions, v_i , are bound between 0 (worst outcome) and a best outcome (e.g., 1). They can be valued using a variety of techniques, including using a direct rating scale (Keeney 1992). The importance of criteria i is represented through swing weights, where the weight, w_i , represents the scale and relative importance of the i^{th} criterion (Diaby and Goeree 2014; Goodwin and Wright 2010; Belton and Stewart 2002).

The aggregate form of the value function is the simplest application, but relies on several assumptions to justify additive aggregation. The first assumption is first-order preference independence, which states that decisions can be made on a subset of criteria irrespective of the other criteria (Keeney and Raiffa 1993). Suppose there are two alternatives under consideration, and they are different on $r < m$ criteria. Define D to be the set of criteria on which the alternatives differ. Assume the criteria that are not in set D are held constant (i.e., they are identical between the alternatives). By definition, the partial value functions are equal for the criteria not in set D . As such program a is preferred to program b if and only if

$$\sum_{i \in D} w_i v_i(a) > \sum_{i \in D} w_i v_i(b) \quad (2.2)$$

This implies that decision-makers can have meaningful preference orderings on a set of criteria without considering the levels of performance on other criteria, provided that the other criteria remain fixed.

The second assumption is that the partial value function is on an interval scale. The interval scale assumption dictates that equal increments on a partial value function represent equal quantitative distances within the criterion (e.g., on a scale between 1 and 10, the value of the difference between 1 and 2 is the same as the difference between 8 and 9). In this way, interval scales provide information about order and possess equal trade-offs across equal intervals within the scale.

The final assumption is the trade-off condition, which satisfies the notion that the weights are scaling constants that render the value scales commensurate (Keeney and Raiffa 1993; Belton and Stewart 2002). This condition is achieved through swing weights, which represent the gain in value by going from the worst value to the best value in each criterion. For example, suppose two partial value functions for two criteria are constructed. Next, suppose that program alternatives a and b differ in terms of two criteria, r and s , which are equally preferred. As such $V(a) = V(b)$. This implies that $w_r v_r(a) + w_s v_s(a) = w_r v_r(b) + w_s v_s(b)$. For this equality to hold, simple algebra demonstrates that the weights are required such that $w_r/w_s = v_s(a) - v_s(b) / v_s(a) - v_s(b)$.

2.4.1.2 Multi-Attribute Utility Theory (MAUT)

MAUT is an extension of von Neumann-Morgenstern (VNM) expected utility theory (Von Neumann and Morgenstern 1953) because it incorporates multi-attribute alternatives (Keeney and Raiffa 1993). The theoretical underpinnings of VNM are similar to value measurement, but MAUT crucially allows for preference relations that involve uncertain outcomes of a particular course of action, which involves risk typically represented through lotteries. To accommodate the idea of risky choices, preferences between lotteries are incorporated directly into the assumptions of preference relations. The first VNM axiom is preferences exist and are transitive, i.e., risky alternatives can be compared such that either a is preferred to b , or b is preferred to c , or they are equally preferred; and if a is preferred to b and b is preferred to c , then a is preferred to c . The second axiom is independence, which states that the preference ordering between two risky goods is independent of the inclusion of a third risky good. To illustrate, suppose in the first stage there are three risky goods (a, b, c), where risky good a is associated with probability p_1 and risky good b with probability $(1-p_1)$. The independence axioms suggest that if $a \geq b$, then $p_1a + (1-p_1)c \geq p_1b + (1-p_1)c$. The axiom assumes that if a decision-maker is comparing good a to good b , their preference should be independent of probability p_1 and good c . This is also called the independence of irrelevant alternatives assumption because if risky good c is substituted for part of a and part of b , it should not change the rankings. The third axiom, continuity, is a mathematical assumption that preferences are continuous. It states that if there are three outcomes such that outcome z_i is preferred to z_j and z_j is preferred to z_k , there is a probability p_1 at which the individual is indifferent between outcome z_j with certainty or receiving the risky prospect made of outcome z_i with probability p_1 and outcome z_k with probability $1-p_1$.

These axioms guarantee the existence of a real-valued utility function such that a is preferred to b if and only if the expected utility of a is greater than the expected utility of b (Mas-Colell et al. 1995). The axiom of continuity provides the guide to making decisions: choose the course of action associated with the greatest sum of probability-weighted utilities. This is the expected utility rule. To apply this rule, the probability and utility associated with possible consequence of each course of an action needs to be assessed, and the probability and utility are multiplied together for each consequence (Fig. 2.3). The products are then added together to obtain the expected utility, $U(z_i)$, of an alternative course action. The process is repeated for each course of action, and the program with the largest expected utility is chosen.

The characterization of VNM utility into a multi-criteria problem depends on utility functions for multiple criteria, $u_i(z_i)$, for $i=1\dots m$ being aggregated into a multi-attribute utility function, $U(z_i)$, that is consistent with lotteries over the criteria. The most common form of the multi-attribute utility function is additive:

$$U(z) = \sum_{i=1}^m k_i u_i(z_i) \quad (2.3)$$

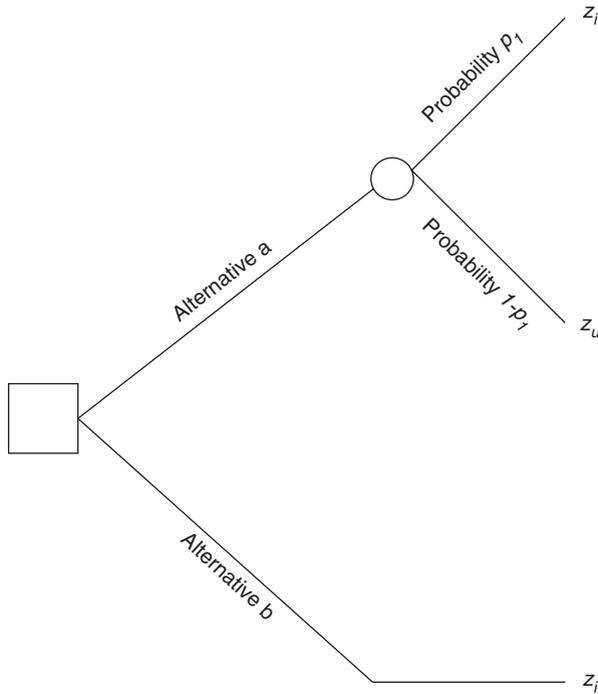


Fig. 2.3 Expected utility theory using the standard gamble method

where k_i is a scaling constant such that $\sum_{i=2}^m k_i = 1$ (Drummond 2005). There are two additional assumptions that are needed to guarantee the existence of an additive utility function: utility independence and additive independence. Utility independence among criteria occurs when there is no interaction between preferences over lotteries among one criterion and the fixed levels for the other criterion; that is, the relative scaling of levels for a given criterion level is constant within each criterion. Additive independence is stronger and suggests that there is no interaction for preference among attributes at all. As such, preference depends on the criterion levels and do not depend on the manner in which the levels of different attributes are combined. The restrictive assumption regarding additive independence can be eased, leading to multiplicative or multi-linear forms of the multi-attribute utility function (Keeney and Raiffa 1993; Drummond 2005).

2.4.1.3 Analytic Hierarchy Process

The analytic hierarchy process (AHP) is a theory of measurement based on mathematics and psychology (Mirkin and Fishburn 1979; Saaty 1980, 1994). It has three principles: (i) decomposition, where a decision problem is structured into a

cluster of hierarchies; (ii) comparative judgments, where judgments of preference are made between criterion attribute levels in each cluster; and (iii) synthesis of priorities, which is the aggregation model. Described below are the four axioms that support AHP: the reciprocal property, homogeneity within criterion, the synthesis axiom, and the expectation axioms (Saaty 1986). The AHP axioms allow for the derivation of ratio scales of absolute numbers through decision-makers’ responses to pairwise comparisons.

The reciprocal axiom requires that a paired comparison between criterion attribute levels, x_1 and x_2 and denoted as $P(x_1, x_2)$, exhibits a property such that if criterion level x_1 is preferred to x_2 and S represents the strength of preference, then the comparison of x_2 with x_1 is $P_c(x_2, x_1) = 1/S$ (Belton and Stewart 2002; Saaty 1980). That is, if x_1 is preferred twice as much as x_2 , then x_2 is preferred one-half as much as x_1 . The homogeneity axiom requires that preferences for attribute criterion levels should not differ too much in terms of strength of preference. For example, questions of a general form such as “How important is criteria level x_1 relative to x_2 ?” are asked (Fig. 2.4). The numerical value of importance is captured from categories that describe strength of preference and are within an order of magnitude from equally important (index 1) to extremely more important (index of 9). The synthesis axiom requires that preferences regarding criterion in a higher-level hierarchy are independent on lower-level elements in the hierarchy (Saaty 1986). Finally, the expectation axiom states that the outcome of the AHP exercise is one such that decision-makers’ preferences or expectations are adequately represented by the outcomes of the exercise (Saaty 1986).

| How important is x_1 relative to x_2 ? | Preference index assigned | Description |
|--|---------------------------|---|
| Equally preferred | 1 | Two criterion levels contribute equally |
| Moderately preferred | 3 | Experience and judgment moderately favour one level over another |
| Strongly preferred | 5 | Experience and judgment strongly favour one level over another |
| Very strongly preferred | 7 | A criterion level is strongly favoured and its dominance demonstrated in practice |
| Absolutely preferred | 9 | The evidence favouring one criterion level over another is of the highest possible order of affirmation |
| Intermediate values are assigned between adjacent categories | 2,4,6,8 | When compromise is needed |

Fig. 2.4 Analytic hierarchy process preference index based on Saaty (1980)

The analysis of AHP paired comparisons judgment data includes the utilization of a comparison matrix. Elements on the principal diagonal of the matrix are equal to 1 because each criterion is at least as good as itself. The off-diagonal elements of the matrix are not symmetric and represent the numerical scale of preference strength expressed as a ratio (as required by the first axiom). The formal analysis requires that the set of value estimates, $v_i(a)$, be consistent with the relative values expressed in the comparison matrix. While there is complete consistency in the reciprocal judgments for any pair, consistency of judgments between alternatives is not guaranteed (Belton and Stewart 2002). The task is to search for the v_i that will provide a fit to the observations recorded in the comparison matrix. This is accomplished through the eigenvector corresponding with the maximum eigenvalue for the matrix. An alternative approach is to calculate the geometric mean of each row of the comparison matrix; each row corresponds to the weight for a specific criterion (Dijkstra 2011).

In the aggregation model for AHP, the importance weight, w_i , of the parent criterion in the hierarchy needs to be calculated. To do this, the above process of pairwise comparisons is applied, where the comparison is between the parent criteria in the hierarchy level. The final aggregation model that allows decision-makers to rank alternatives is similar to the value measurement approach because an additive aggregation is used:

$$P(a) = \sum_{i=1}^m w_i v_i(a) \quad (2.4)$$

where $P(a)$ is the priority of alternative a , $v_i(a)$ is the partial value function of a criterion level, and w_i is the overall weight of criterion.

2.4.2 Outranking

Outranking utilizes the concept of dominance between partial preference functions of the alternatives (Belton and Stewart 2002). In MCDA, dominance is defined as $z_i(a) \geq z_i(b)$ for all criteria i , where there is strict inequality on at least one criterion j , i.e., $z_j(a) > z_j(b)$ (Belton and Stewart 2002). Dominance rarely occurs in real-world decision-making. Outranking generalizes the definition of dominance by defining an outranking relation that represents a binary condition on a set of alternatives, denoted by A , such that program a will outrank b if there is evidence suggesting that “program a is at least as good as program b .” The outranking relation is represented by aSb for $(a, b) \in A$ (Ehrgott et al. 2010). Of note, outranking investigates the hypothesis that aSb by focusing on whether there is compelling evidence for the hypothesis (i.e., “strong enough”), rather than focusing on strength of preference using compensatory preference structures (Belton and Stewart 2002). As a result, in addition to the possibility of dominance or indifference, there may be a lack of compelling evidence to conclude dominance or indifference (Belton and Stewart 2002).

The outranking relation is constructed using concordance and discordance indices. These indices characterize the sufficiency of evidence for or against one alternative outranking another. The concepts of concordance and discordance can be stated as follows:

- *Concordance*: For an outranking aSb to be true, there must be a sufficient majority of criterion to favor the assertion.
- *Non-discordance*: When concordance holds, none of the criteria in the minority should oppose too strongly the assertion that aSb ; alternatively, *discordance* is where b is very strongly preferred to a on one or more of the minority of criteria which call into question the hypothesis that aSb (Figueira et al. 2005).

The process of characterizing concordance and discordance starts with evaluating the performance of each alternative on the criterion using a decision matrix summarizing the partial preference functions. The matrix is structured such that each row summarizes the partial preference function for the individual criterion which is located in each of the m columns. Outranking recognizes that partial preference functions are not precise (Belton and Stewart 2002). Indifference thresholds, defined as $p_i[z]$ and $q_i[z]$, are used to acknowledge a distinction between weak and strict preference, where alternative a is weakly preferred to alternative b for criterion i if $z_i(a) > z_i(b) + q_i[z_i(b)]$; it follows that $z_i(a) - z_i(b) > q_i[z_i(b)]$. Alternative a is strictly preferred to alternative b for criterion i if $z_i(a) > z_i(b) + p_i[z_i(b)]$ and $z_i(a) - z_i(b) > p_i[z_i(b)]$. In this notation, it would be necessary for $q_i[z_i(b)] > p_i[z_i(b)]$ to distinguish between weak and strict thresholds (Belton and Stewart 2002). Indifference between a and b can happen when there is not strict inequality between the partial preference function. While the decision matrix demonstrates if alternative a outperforms alternative b in each of the criterion, it does not account for the relative importance of the criteria. This is achieved through criterion weights (w_i). The weights measure the influence that each criterion should have in building the case for one alternative or another.

There are several approaches to estimating concordance or discordance, including the ELECTRE methods (Roy 1991), PROMETHEE (Brans and Vincke 1985), and GAIA (Brans and Mareschal 1994). We present ELECTRE1 below. The concordance index, $C(a,b)$, characterizes the strength of support for the hypothesis that program a is at least as good as program b . The discordance index, $D(a,b)$, measures the strength of evidence against aSb . In ELECTRE1, the concordance index is:

$$C(a,b) = \frac{\sum_{j \in Q(a,b)} w_j}{\sum_{i=1}^m w_i} \quad (2.5)$$

where $Q(a,b)$ is the set of criterion for which a is equal or preferred to b as determined by the decision matrix. Note that the concordance index is bound between 0 and 1, and as $C(a,b)$ approaches 1, there is stronger evidence in support

of the claim that a is preferred to b . A value of 1 indicates that program a dominates b on all criterion.

The discordance index will differ depending on if the decision matrix values are cardinal or if the weights are on a scale that is comparable across criteria (Belton and Stewart 2002). When these conditions hold, the discordance index is:

$$D(a,b) = \frac{\max_{i \in R(a,b)} [w_i (z_i(b) - z_i(a))]}{\max_{i=1}^m \max_{c,d \in A} [w_i | (z_i(c) - z_i(d))]} \quad (2.6)$$

where $R(a,b)$ is the set of criteria for which b is strictly preferred to a in the set of A alternatives. The index is calculated as the maximum weighted estimate for which program b is better than program a divided by the maximum weighted difference between any two alternatives on any criterion. Note that in the two alternative cases, an instance of b outperforming a would result in a value of 1. When the partial preference scores are not cardinal (e.g., they are qualitative relations) or when the criteria importance weights are not on a scale that is comparable across criteria, the discordance criterion is based on a veto threshold. That is, for each criterion i , there is a veto threshold defined as t_i such that program a cannot outrank b if the score for b on any criterion exceeds the score for a on that criterion. The discordance index is:

$$D(a,b) = \begin{cases} 1 & \text{if } z_i(b) - z_i(a) > t_i \text{ for any } i \\ 0 & \text{otherwise} \end{cases} \quad (2.7)$$

The concordance and discordance indices are evaluated against thresholds, denoted as C^* and D^* , to determine the relationship between program a and b according to the following: if $C(a,b) > C^*$ and $D(a,b) < D^*$ then aSb , otherwise a does not outrank b ; program bSa if $C(b,a) > C^*$ and $D(b,a) < D^*$, otherwise b does not outrank a . There is indifference between program a and b when aSb and bSa . The two are incomparable if neither program outranks the other, i.e., not aSb and not bSa . A summary of these outranking relations, adapted from Belton and Stewart (2002), is in Fig. 2.5.

2.4.3 Goal Programming

Goal programming attempts to model complex multi-criteria problems using concepts linked to a decision heuristic called “satisficing” (Belton and Stewart 2002). Introduced by Simon (1976), satisficing is a cognitive heuristic where a decision-maker examines the characteristics of multiple alternatives until an acceptability threshold is achieved. This is in contrast to a compensatory

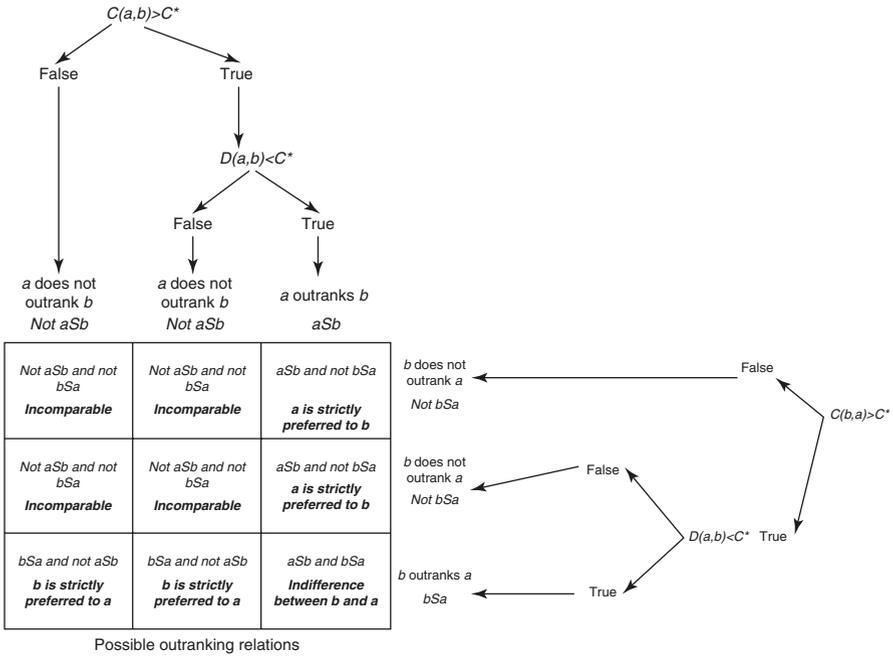


Fig. 2.5 Outranking relations from a pairwise comparison

framework, which assumes decision-makers have unlimited cognitive processing abilities and carefully consider all information(Regier et al. 2014b; Kahneman 2003; Simon 2010). Because of complexity, decision-makers seek a choice that is “good enough,” i.e., a choice that is satisfactory rather than optimal.

Goal programming operationalizes satisficing by assuming decision-makers seek to achieve a satisfactory level of each criterion (Tamiz et al. 1995). Attention is shifted to other criteria when a desired threshold is achieved. Goal programming methods encapsulate the following two assumptions. First, each criterion is associated with an attribute that is measurable and represented on a cardinal scale. Thus, the methods utilize partial preference functions, denoted by $z_i(a)$, where $i=1..m$ criteria for alternative a . Second, decision-makers express judgment in terms of goals or “aspiration levels” for the m criterion, which are understood in terms of desirable levels of performance (e.g., ICER is below a threshold of \$50,000 per QALY). In goal programming notation, goals are denoted by g_i where $i=1, \dots, m$. With the goals defined, an algorithm is used identify the alternatives which satisfy the goals in an order of priority (Thokala and Duenas 2012; Ignizio 1978).

Decision-makers’ preferences for goals will differ depending on the context or frame of each criterion. That is, the direction of preference as reflected in each goal will differ depending if the attribute criterion, $z_i(a)$, is defined in the context of: maximization, with the goal representing a minimum level of satisfactory performance; or minimization of $z_i(a)$, with the goal of representing the maximum

level of tolerance; or whether the goal is to reach a desirable level of performance for $z_i(a)$ (Belton and Stewart 2002). The difference between $z_i(a)$ and g_i is denoted by δ_i^- or δ_i^+ , which respectively represents the quantitative amount the partial preference function is under- or overachieved.

The solution to the decision-makers problem of satisficing is investigated through mathematical optimization techniques (e.g., linear programming) that aim to achieve the best outcome given an objective that is subject to constraints (e.g., minimize the deviation of attribute values subject to a value function). Goal programming models can be categorized either through a weighted goal programming or one that focuses on lexicographic preferences. In weighted goal programming, deviations from goals are minimized after importance weights are assigned to each of the $z_i(a)$. This can be achieved through following algebraic formulation (Tamiz et al. 1998; Rifai 1996; Kwak and Schniiederjans 1982):

$$\min Z = \sum_{i=1}^m (w_i^- \delta_i^- + w_i^+ \delta_i^+) \quad (2.8)$$

$$\text{subject to } f_i(x) + \delta_i^- - \delta_i^+ = g_i \text{ for } i = 1, \dots, m$$

where x is a vector of the decision variables that can be varied such that a criterion achieves a specified goal (e.g., the price of a drug), $f_i(x)$ is a linear objective function equivalent to the realization of the partial preference value $z_i(a)$ for an x vector, g_i is the target value for each $z_i(a)$, δ_i^- and δ_i^+ are the negative and positive deviation from the target values, and w_i^- , w_i^+ are the importance weights. Of note, the importance weights need to conform to a trade-off condition when the weighted sum approach is used (e.g., through swing weights).

In lexicographic goal programming, deviational variables are assigned into priority levels and minimized in a lexicographic order (Belton and Stewart 2002). A lexicographic ordering is one where a decision-maker prefers any amount of one criterion to any amount of the other criteria. Only when there is a tie does the decision-maker consider the next most preferred criterion. In goal programming, a sequential lexicographic minimization of each priority criterion is conducted while maintaining the minimal values reached by all higher priority level minimizations (Ijiri 1965):

$$\text{Lex min } O = (g_1(\delta^+, \delta^-), g_2(\delta^+, \delta^-), \dots, g_m(\delta^+, \delta^-)) \quad (2.9)$$

$$\text{subject to } f_i(x) + \delta_i^- - \delta_i^+ = g_i \text{ for } i = 1, \dots, m$$

where all definitions above save O , which is an ordered vector of priorities (Tamiz et al. 1995). From a practical perspective, priority classes on each criterion are defined, and minimization of the weighted sum is conducted in relation to the goal.

Once the solution is obtained for the higher-order priority, the second priority class is minimized subject to the constraint that the weighted sum from goals in the first priority class does not exceed what was determined in the first step. The process is continued through each priority class in term.

2.5 Concluding Remarks

This chapter introduced the theoretical foundations and methods that support MCDA. MCDA provides decision-makers with a set of tools that can aid stakeholders in making consistent and transparent decisions. MCDA methods draw on theories that account for both the qualitative and quantitative aspects of decision-making. This is achieved through a process that includes a comprehensive approach to problem structuring and model building. We conclude by noting that there is a paucity of MCDA theory-related research in healthcare. We encourage future research that explores which MCDA methods best address stakeholders' needs in the context of the unique challenges we face in improving health and healthcare.

Acknowledgments The Canadian Centre for Applied Research in Cancer Control (ARCC) is funded by the Canadian Cancer Society Research Institute (ARCC is funded by the Canadian Cancer Society Research Institute grant #2015-703549). Dr. Stuart Peacock is supported by the Leslie Diamond Chair in Cancer Survivorship, Faculty of Health Sciences, Simon Fraser University.

References

- Ackoff RL (1979a) Future of operational-research is past. *J Oper Res Soc* 30(2):93–104. doi:[10.2307/3009290](https://doi.org/10.2307/3009290)
- Ackoff RL (1979b) Resurrecting the future of operational-research. *J Oper Res Soc* 30(3): 189–199. doi:[10.2307/3009600](https://doi.org/10.2307/3009600)
- Arnstein SR (1969) Ladder of citizen participation. *J Am Inst Plann* 35(4):216–224. doi:[10.1080/01944366908977225](https://doi.org/10.1080/01944366908977225)
- Baltussen R, Niessen L (2006) Priority setting of health interventions: the need for multi-criteria decision analysis. *Cost Effectiveness and Resource Allocation : C/E* 4:14. doi:[10.1186/1478-7547-4-14](https://doi.org/10.1186/1478-7547-4-14)
- Baltussen R, Stolk E, Chisholm D, Aikins M (2006) Towards a multi-criteria approach for priority setting: an application to Ghana. *Health Econ* 15(7):689–696. doi:[10.1002/hec.1092](https://doi.org/10.1002/hec.1092)
- Belton V, Stewart T (2010) Problem structuring and multiple criteria decision analysis. In: *Trends in multiple criteria decision analysis*, vol 142, International series in operations research & management science. Springer, New York, p xvi, 412 p
- Belton V, Stewart TJ (2002) *Multiple criteria decision analysis: an integrated approach*. Kluwer Academic Publishers, Boston
- Brans JP, Mareschal B (1994) The Promcalc and Gaia decision-support system for multicriteria decision aid. *Decis Support Syst* 12(4–5):297–310. doi:[10.1016/0167-9236\(94\)90048-5](https://doi.org/10.1016/0167-9236(94)90048-5)
- Brans JP, Vincke PH (1985) A preference ranking organization method – (the Promethee method for multiple criteria decision-making). *Manag Sci* 31(6):647–656. doi:[10.1287/mnsc.31.6.647](https://doi.org/10.1287/mnsc.31.6.647)

- Camerer C (1995) Individual decision making. In: Kagel J, Roth A (ed) *The Handbook of Experimental Economics*. Princeton University Press, NJ
- Checkland P (1981) *Systems thinking, systems practice*. J. Wiley, Chichester/New York
- Devlin N, Sussex J (2011) *Incorporating multiple criteria in HTA: methods and processes*. Office of Health Economics, London
- Diaby V, Goeree R (2014) How to use multi-criteria decision analysis methods for reimbursement decision-making in healthcare: a step-by-step guide. *Expert Rev Pharmacoecon Outcomes Res* 14(1):81–99. doi:[10.1586/14737167.2014.859525](https://doi.org/10.1586/14737167.2014.859525)
- Dijkstra T (2011) On the extraction of weight from pairwise comparison matrices. *CEJOR* 21(1):103–123
- Dolan JG (2010) Multi-criteria clinical decision support: a primer on the use of multiple criteria decision making methods to promote evidence-based, patient-centered healthcare. *Patient* 3(4):229–248. doi:[10.2165/11539470-000000000-00000](https://doi.org/10.2165/11539470-000000000-00000)
- Drummond M (2005) *Methods for the economic evaluation of health care programmes*, 3rd edn, Oxford medical publications. Oxford University Press, Oxford England/New York
- Ehrgott M, Figueira J, Greco S (2010) Trends in multiple criteria decision analysis, vol 142, *International series in operations research & management science*. Springer, New York
- Figueira J, Greco S, Ehrgott M (2005) *Multiple criteria decision analysis : state of the art surveys*. Springer, New York
- Goetghebeur MM, Wagner M, Khoury H, Levitt RJ, Erickson LJ, Rindress D (2008) Evidence and value: impact on DEcisionMaking--the EVIDEM framework and potential applications. *BMC Health Serv Res* 8:270. doi:[10.1186/1472-6963-8-270](https://doi.org/10.1186/1472-6963-8-270)
- Goodwin P, Wright G (2010) *Decision analysis for management judgment*, 4th edn. Wiley, Hoboken
- Hester PT, Adams KM (2014) Problems and messes. 26:23–34. doi:[10.1007/978-3-319-07629-4_2](https://doi.org/10.1007/978-3-319-07629-4_2)
- Ignizio J (1978) A review of goal programming: a tool for multiobjective analysis. *J Oper Res Soc* 29:1109–1119
- Ijiri Y (1965) *Management goals and accounting for control*. Studies in mathematical and managerial economics, vol 3. North Holland Pub. Co., Amsterdam
- Kahneman D (2003) A perspective on judgment and choice – mapping bounded rationality. *Am Psychol* 58(9):697–720. doi:[10.1037/0003-066X.58.9.697](https://doi.org/10.1037/0003-066X.58.9.697)
- Keeney RL (1992) *Value-focused thinking: a path to creative decision making*. Harvard University Press, Cambridge
- Keeney RL, Mcdaniels TL (1992) Value-focused thinking about strategic decisions at Bc Hydro. *Interfaces* 22(6):94–109. doi:[10.1287/Inte.22.6.94](https://doi.org/10.1287/Inte.22.6.94)
- Keeney RL, Raiffa H (1993) *Decisions with multiple objectives : preferences and value tradeoffs*. Cambridge University Press, Cambridge/New York
- Kimball AW (1957) Errors of the 3rd kind in statistical consulting. *J Am Stat Assoc* 52(278):133–142. doi:[10.2307/2280840](https://doi.org/10.2307/2280840)
- Kwak NK, Schniederjans MJ (1982) An alternative method for solving goal programming-problems – a reply. *J Oper Res Soc* 33(9):859–860
- Lancaster KJ (1966) New approach to consumer theory. *J Polit Econ* 74(2):132–157. doi:[10.1086/259131](https://doi.org/10.1086/259131)
- MacLean S, Burgess MM (2010) In the public interest: assessing expert and stakeholder influence in public deliberation about biobanks. *Public Underst Sci* 19(4):486–496. doi:[10.1177/0963662509335410](https://doi.org/10.1177/0963662509335410)
- Mas-Colell A, Whinston MD, Green JR (1995) *Microeconomic theory*. Oxford University Press, New York
- Mirkin BG, Fishburn PC (1979) *Group choice*. Scripta series in mathematics. V. H. Winston; distributed by Halsted Press, Washington, D.C.
- Montibeller G, von Winterfeldt D (2015) Cognitive and motivational biases in decision and risk analysis. *Risk Anal* 35(7):1230–1251. doi:[10.1111/risa.12360](https://doi.org/10.1111/risa.12360)

- NICE (2008) Guide to the Methods of Technology Appraisals. National Institute of Clinical Excellence, London, England
- O'Doherty KC, Hawkins AK, Burgess MM (2012) Involving citizens in the ethics of biobank research: informing institutional policy through structured public deliberation. *Soc Sci Med* 75(9):1604–1611. doi:[10.1016/j.socscimed.2012.06.026](https://doi.org/10.1016/j.socscimed.2012.06.026)
- Peacock S, Mitton C, Bate A, McCoy B, Donaldson C (2009) Overcoming barriers to priority setting using interdisciplinary methods. *Health Policy* 92(2–3):124–132. doi:[10.1016/j.healthpol.2009.02.006](https://doi.org/10.1016/j.healthpol.2009.02.006)
- Phillips LD (1984) A theory of requisite decision-models. *Acta Psychol* 56(1–3):29–48. doi:[10.1016/0001-6918\(84\)90005-2](https://doi.org/10.1016/0001-6918(84)90005-2)
- Phillips LD, Phillips MC (1993) Facilitated work groups – theory and practice. *J Oper Res Soc* 44(6):533–549. doi:[10.1057/Jors.1993.96](https://doi.org/10.1057/Jors.1993.96)
- Rabin M (1998) Psychology and economics. *J Econ Lit* 36(1):11–46
- Rawlins M, Barnett D, Stevens A (2010) Pharmacoeconomics: NICE's approach to decision-making. *Br J Clin Pharmacol* 70(3):346–349. doi:[10.1111/j.1365-2125.2009.03589.x](https://doi.org/10.1111/j.1365-2125.2009.03589.x)
- Regier DA, Bentley C, Mitton C, Bryan S, Burgess MM, Chesney E, Coldman A, Gibson J, Hoch J, Rahman S, Sabharwal M, Sawka C, Schuckel V, Peacock SJ (2014a) Public engagement in priority-setting: results from a pan-Canadian survey of decision-makers in cancer control. *Soc Sci Med* 122:130–139. doi:[10.1016/j.socscimed.2014.10.038](https://doi.org/10.1016/j.socscimed.2014.10.038)
- Regier DA, Watson V, Burnett H, Ungar WJ (2014b) Task complexity and response certainty in discrete choice experiments: an application to drug treatments for juvenile idiopathic arthritis. *J Behav Exp Econ* 50:40–49. doi:[10.1016/j.socec.2014.02.009](https://doi.org/10.1016/j.socec.2014.02.009)
- Rifai AK (1996) A note on the structure of the goal-programming model: assessment and evaluation. *Int J Oper Prod Manag* 16(1):40. doi:[10.1108/01443579610106355](https://doi.org/10.1108/01443579610106355)
- Rowe G, Frewer LJ (2000) Public participation methods: a framework for evaluation. *Sci Technol Hum Values* 25(1):3–29
- Roy B (1991) The outranking approach and the foundations of electre methods. *Theor Decis* 31(1):49–73. doi:[10.1007/Bf00134132](https://doi.org/10.1007/Bf00134132)
- Roy B (1996) Multicriteria methodology for decision aiding, vol 12, Nonconvex optimization and its applications. Kluwer Academic Publishers, Dordrecht/Boston
- Saaty TL (1980) The analytic hierarchy process : planning, priority setting, resource allocation. McGraw-Hill International Book Co, New York/London
- Saaty TL (1986) Axiomatic foundation of the analytic hierarchy process. *Manag Sci* 32(7):841–855. doi:[10.1287/mnsc.32.7.841](https://doi.org/10.1287/mnsc.32.7.841)
- Saaty TL (1994) Fundamentals of decision making and priority theory with the analytic hierarchy process, vol 6, 1st edn, Analytic hierarchy process series. RWS Publications, Pittsburgh
- Schoner B, Choo E, Wedley W (1999) Comment on 'Rethinking value elicitation for personal consequential decision' by G wright and P goodwin. *J Multi-Criteria Decis Anal* 8:24–26
- Simon HA (1976) Administrative behavior: a study of decision-making processes in administrative organization, 3rd edn. Free Press, New York
- Simon HA (2010) A behavioral model of rational choice. *Compet Policy Int* 6(1):241–258
- Tamiz M, Jones D, Romero C (1998) Goal programming for decision making: an overview of the current state-of-the-art. *Eur J Oper Res* 111(3):569–581. doi:[10.1016/S0377-2217\(97\)00317-2](https://doi.org/10.1016/S0377-2217(97)00317-2)
- Tamiz M, Jones DE, Eldarzi E (1995) A review of goal programming and its applications. *Ann Oper Res* 58:39–53. doi:[10.1007/Bf02032309](https://doi.org/10.1007/Bf02032309)
- Thokala P, Duenas A (2012) Multiple criteria decision analysis for health technology assessment. *Value Health : J Int Soc Pharmacoeconomics Outcomes Res* 15(8):1172–1181. doi:[10.1016/j.jval.2012.06.015](https://doi.org/10.1016/j.jval.2012.06.015)
- Von Neumann J, Morgenstern O (1953) Theory of games and economic behavior, 3rd edn. Princeton University Press, Princeton