



Published in final edited form as:

Patient. 2010 ; 3(4): 229–248. doi:10.2165/11539470-000000000-00000.

Multi-criteria clinical decision support: A primer on the use of multiple criteria decision making methods to promote evidence-based, patient-centered healthcare

James G. Dolan, MD

Department of Community and Preventive Medicine, University of Rochester, Rochester, New York, USA

Abstract

Current models of healthcare quality recommend that patient management decisions be evidence-based and patient-centered. Evidence-based decisions require a thorough understanding of current information regarding the natural history of disease and the anticipated outcomes of different management options. Patient-centered decisions incorporate patient preferences, values, and unique personal circumstances into the decision making process and actively involve both patients along with health care providers as much as possible. Fundamentally, therefore, evidence-based, patient-centered decisions are multi-dimensional and typically involve multiple decision makers.

Advances in the decision sciences have led to the development of a number of multiple criteria decision making methods. These multi-criteria methods are designed to help people make better choices when faced with complex decisions involving several dimensions. They are especially helpful when there is a need to combine “hard data” with subjective preferences, to make trade-offs between desired outcomes, and to involve multiple decision makers. Evidence-based, patient-centered clinical decision making has all of these characteristics. This close match suggests that clinical decision support systems based on multi-criteria decision making techniques have the potential to enable patients and providers to carry out the tasks required to implement evidence-based, patient-centered care effectively and efficiently in clinical settings.

The goal of this paper is to give readers a general introduction to the range of multi-criteria methods available and show how they could be used to support clinical decision-making. Methods discussed include the balance sheet, the even swap method, ordinal ranking methods, direct weighting methods, multi-attribute decision analysis, and the analytic hierarchy process (AHP)

Introduction

Imagine you are a health care provider seeing Mrs. Gray, a married, 50 year old mother of three who owns a small art supply store. You are meeting to discuss treatment of a newly diagnosed, as yet untreated chronic disease that is causing symptoms severe enough to force her to reduce the time she spends at her shop. You are both aware that the disease is likely to progress and more severely interfere with her ability to manage her business, as well as her other daily activities, in the future. You both agree that some form of active treatment should be started.

Fortunately, several treatment options are available. Drug A has been used for many years. It effectively alleviates symptoms, halts disease progression in two-thirds of patients, and is very well tolerated. Two newer drugs, B₁ and B₂, are also available. They are slightly less effective than Drug A in relieving symptoms but more effective in halting disease progression. Both drugs can cause a rare but serious side effect. Finally, there is a newly released drug, Drug C, which is reported to both alleviate symptoms and halt disease progression more effectively than any of the other available medications. However, it has higher risks of both serious and common side effects, and, because it is a new medication, could also potentially cause additional, currently unknown, adverse effects. Mrs. Gray's monthly out-of-pocket expense for Drug A would be relatively low. Her costs for the B drugs would be three to six times higher. Drug C will cost her 10 times more. Which medication should you prescribe?

Current models of healthcare quality recommend that this decision be evidence-based and patient-centered. [1–2] Evidence-based decisions are based on a thorough understanding of current information regarding the natural history of the disease and the anticipated outcomes of alternative courses of management. Patient-centered decisions actively involve patients as much as possible and incorporate patient preferences, values, and their unique circumstances in the decision-making process.

In many contemporary clinical settings, however, you and Mrs. Gray would find it difficult to make this decision in a patient-centered, evidence-based manner. Specific information about the range of outcomes she is likely to experience after starting any of the medications will probably be hard to find. [3] Even if it is available, your busy practice schedule will severely limit the time you and Mrs. Gray have to review the advantages and disadvantages of the different drugs and make the trade-offs needed to select which one to use. In addition to these practical concerns, you and Mrs. Gray might find it difficult to make a truly good decision. When an ideal solution is not available, it is always unsettling to make trade-offs between the advantages and disadvantages of several imperfect alternatives. It is even more disconcerting when the actual outcomes that will follow any decision are uncertain. Finally, there is no common language you can use to accurately and easily discuss your decision-related preferences with each other.

Many common clinical decisions involve similarly complex choices. These difficulties in making evidence-based, patient centered decisions in busy clinical settings are well recognized. One of the proposed solutions is the development and implementation of new clinical decision support systems. [1,4–5] To be successful, they will have to help clinicians, patients, and other involved stakeholders make better choices when faced with complex decisions that involve trade-offs between the pros and cons of imperfect options, a mix of objective data and subjective judgments, and uncertain future outcomes.

Decisions with these characteristics are not unique to health care; they commonly occur in many areas of human endeavor. To help people make better choices when faced with such a decision, a number of multi-criteria decision making methods have been developed. [6–7] Multi-criteria methods are a type of decision analysis specifically designed for use in situations where it is important to transparently incorporate multiple considerations into a decision making process. Their main goal is to help decision makers make better choices by helping them achieve greater understanding and insight into the decision they are facing.

Interest in multi-criteria decision making methods has been growing rapidly. Between 2000 and 2006, the number of publications about multi-criteria decision making grew exponentially from just over 1,500 to nearly 6,500. The leading application areas over this time have been management science, operations research, computer science, environmental

science, engineering, economics, energy, and water sources. [8] Although small in comparison to other areas, there has also been a growing interest in the use of multi-criteria methods in healthcare decision making. [9–13]

Because multi-criteria methods are designed to support decision making in complex circumstances identical to those posed by many common patient management decisions, they represent a potentially useful foundation on which to build new clinical decision support systems to facilitate the provision of evidence-based, patient-centered care. The goal of this paper is to introduce readers to a sample of multi-criteria methods and illustrate how they could be used to support clinical decision making.

Multi-criteria decision making methods

For this paper, I will broadly define a multi-criteria decision method as one that guides the user through an evaluation of potential decision options using explicit profiles of their advantages and disadvantages across a range of distinct dimensions. This process of breaking a decision down into smaller units for analysis is called decomposition. Several terms are used for these dimensions depending on the perspective taken. If based on the options, they are called attributes or characteristics. If based on the goal of the decision they are called objectives or criteria. To be consistent with the term most commonly used to describe the field as a whole, I will use the term criteria in this paper. A glossary of these and other terms used to describe and implement multi-criteria decision making methods is included in Appendix 1.

Multi-criteria methods can be categorized in several ways. When classified based on how the decision criteria are used, methods can be either compensatory or non-compensatory. Compensatory methods incorporate information from all the decision criteria whereas non-compensatory methods do not. Examples of both types are illustrated below.

Multi-criteria methods can also be described based on the analytic strategy used. Value-based approaches develop quantitative measures of how well the options fulfill the criteria and the relative priorities of the criteria in achieving the goal of the decision. These measures can be created in a variety of ways which accounts for most of the differences among the methods included in this category. Value-based methods are the most widely used type of multi-criteria method and most of the methods discussed here are in this category. Other strategy-based categories include the even swap method, outranking methods, and goal-based methods. The even swap method is discussed below. Detailed descriptions of the other methods are beyond the scope of this paper and can be found in recent reviews. [6–7]

Conjoint analysis, also called discrete choice analysis, is another method that is frequently used to address multi-criteria decision problems. Conjoint analysis derives preference weights for the attributes of decision options based on a series of choices people make between hypothetical alternatives that contain different combinations of attribute levels. A fundamental difference between conjoint analysis and multi-criteria methods is how the preference scores are derived. Conjoint analysis uses an indirect approach; multi-criteria methods take a direct approach. Conjoint analysis is not usually considered a multi-criteria method and will not be discussed further. Several recent papers discussing the use of conjoint analysis in healthcare decision making are available. [14–18]

Balance sheets

Balance sheet overview—The most basic multi-criteria decision making method is a balance sheet, a table that summarizes information about the alternatives categorized into a

set of mutually exclusive dimensions. Decision support cards and dashboards are closely related methods that present the same information using separate displays for each dimension rather than a single combined table. [19–21]

Table I shows a balance sheet that summarizes the available information about Mrs. Gray's treatment decision. The columns list information describing the attributes of each alternative grouped into three categories:

- ◆ Effectiveness, further divided into controlling symptoms and halting disease progression.
- ◆ Risk of adverse effects, further divided into serious and common side effects. Serious side effects require stopping the medication, additional active treatment, and could result in permanent health consequences. Aplastic anemia is an example. Common side effects primarily cause minor symptoms and may or may not require a change in medication. Examples include rash, headaches, and dyspepsia.
- ◆ Out-of-pocket cost, how much Mrs. Gray would have to pay for the medicine every month.

Clinical decision support using a balance sheet consists of making it readily available to patients and providers and helping them use the information to effectively evaluate the pros and cons of each option. The balance sheet shown in Table I illustrates two ways how this could be done.

First, both data and relative rankings of the alternatives are included. The rankings serve two functions. They clearly show whether high or low values are preferred and they speed the interpretation of the raw data by providing an additional level of organization to the information provided.

The second approach is the identification of a dominated alternative, in this case Drug B₂. Dominated alternatives have no unique advantages and are inferior to at least one other alternative on every dimension. In the example, Drug B₂ is dominated because it is never the best alternative and is inferior to Drugs A and B₁ in terms of symptom relief, common adverse effects, and out-of-pocket cost. In general, dominated alternatives should be eliminated from further consideration so that decision makers can focus their attention on more promising alternatives. A persistent reluctance by decision makers to eliminate dominated alternatives can indicate that an important consideration is missing from the balance sheet. If so, a revised balance sheet should be created.

The balance sheet can also be used to help the decision makers decide if there are any standards alternatives must meet to merit further consideration or if they should simply “take the best”, i.e., choose the alternative that does the best on the most important criterion. [22] For example, Mrs. Gray could either decide to eliminate Drug C from consideration because she is not able to afford it or decide to choose it, regardless of the cost, because it gives her the best chance of preventing progression of her disease.

Decisions like these that are made based on just a subset of the decision criteria are called non-compensatory decisions. They are fast and often useful. Non-compensatory strategies, however, run the risk of eliminating alternatives that represent good choices across a number of criteria because they assume that their combined strengths over the entire range of criteria are not enough to overcome a relative weakness on just one or two dimensions. For this reason, decision makers facing important decisions should use non-compensatory strategies judiciously. In Mrs. Gray's case, the two common non-compensatory strategies described above lead to markedly different decisions: setting a maximum out-of-pocket standard less

than \$100 would eliminate Drug C, the favored alternative if the decision was to “take the best” in terms of effectiveness.

Balance sheet illustration—To illustrate how a balance sheet could be used to support Mrs. Gray’s treatment decision, she and her provider would review the balance sheet illustrated in Table I. Assume that they agree that the dominated option, Drug B₂, can be eliminated from further consideration. (We will continue to make this assumption throughout the rest of the paper.)

Imagine that Mrs. Gray becomes concerned about her ability to afford Drug C. After a brief discussion, however, she and her provider decide to keep it under consideration because the number of alternatives is already manageable and Drug C is the most effective drug. The decision making process then consists of a careful evaluation of the three non-dominated options - Drugs A, B₁, and C - all of which have a unique combination of advantages and disadvantages, culminated by the selection of the one that seems most appropriate.

Balance sheet comments—The use of balance sheets to support medical decision making has been advocated by Eddy. [23] The advantages of this approach stem from its simplicity and familiarity. Many people are familiar with tabular summaries and know how to interpret the information they contain. It is also a practical approach. Assuming an appropriate balance sheet can be created beforehand, it only adds one additional step to the usual care process.

The main disadvantage of the balance sheet approach is that it is unclear whether it provides enough decision support to foster consistent, high quality clinical decision making. [24–25] It is possible that important considerations may not be included on the sheet and therefore omitted from the decision making process. It is also unclear how well balance sheets stimulate assessment of individual preferences regarding decision trade-offs and help the decision makers integrate them into the decision making process.

The even swap method

Even swap overview and illustration—The oldest formally described multi-criteria method is a swapping approach, initially described by Benjamin Franklin in 1762, and later incorporated into the “Even Swap” method. [26–27] This method consists of identifying strengths and weaknesses of the decision alternatives that balance each other out. It is best described using an example.

Since we are assuming that Mrs. Gray and her provider are willing to eliminate Drug B₂ from further consideration, they would start by creating a new balance sheet that contains only the non-dominated alternatives. This simplified balance sheet is shown in Table II.

The even swap technique involves determining changes in the drugs’ profiles that would make them equivalent. For clarity, we will assume that all swaps will be based on Drug A. In actual practice, swaps can involve any alternative.

Imagine that our decision makers decide that they would be willing to trade a \$45 increase in monthly out-of-pocket cost for Drug A, from \$15 to \$50, in return for a 10% absolute increase in effectiveness in halting progression of the disease, from 65% to 75%. This swap yields a revised profile for drug A: the adjusted disease progression effectiveness is now 75% and the adjusted cost is now \$50 per month. By making this swap, Drug A now dominates Drug B₁ and Drug B₁ is eliminated from further consideration.

Now assume they are willing to trade a decrease in symptom relief effectiveness for Drug A from 95% to 85% in return for a 5% increase in effectiveness in halting disease progression from 75% to 80%. They then decide to trade an increase in Drug A's risk of common side effects from 1% to 5% and cost from \$50 to \$75 in return for another 5% absolute increase in effectiveness in halting disease progression. This swap results in a further revised profile for Drug A; its effectiveness in halting disease progression is now 85% and its cost is \$75. This change results in Drug A dominating Drug C indicating it is the most appropriate choice. Table III illustrates these swaps.

Even Swap comments—Compared with the balance sheet approach, the potential advantage of the even swap method is that it could result in a sounder decision making process. The decision is now based on an explicit process of comparing the relative advantages and disadvantages of the decision alternatives. In addition to helping make the judgments more consistent with the decision makers' preferences, this process could also help them discuss the decision more meaningfully.

There are, however, several potential problems with using the even swap method for clinical decision support. First, it is currently unclear whether clinical decision makers are able to reliably make the swaps required to use this approach or if it effectively promotes communication about the decision-making process. Moreover, the input required of the users is complicated and the added time that is required to make the swap-related judgments may make it less feasible for use in busy clinical settings.

A Pubmed search in June 2010 did not reveal any published medical applications of this method. A web-based computer program implementing the smart swap procedure is freely available at <http://www.smart-swaps.hut.fi/>. [28]

Value-based methods

Value-based methods lead the decision maker(s) through a series of judgments that produce quantitative scores that indicate how well they think the alternatives meet the decision criteria and the relative priorities of the decision criteria in making the decision. Methods in this category include ordinal weighting methods, direct weighting, methods based on multi-attribute utility theory (MAUT), and the Analytic Hierarchy Process (AHP).

All of these methods start with the creation of a hierarchical decision model, sometimes called a value tree, that explicitly defines the goal of the decision, the alternatives being considered, and the criteria that will be used to judge how well the alternatives meet the goal. An example is shown in Figure 1. The decision criteria are mutually exclusive characteristics of decision alternatives, identical to the attribute categories used to create balance sheets. The next step is to identify the options that will be considered, summarize information about how well the options meet each criterion, and create a balance sheet. After checking for dominated alternatives, the multi-criteria analysis is performed. This consists of a series of method-specific steps that decision makers follow to generate the scores that indicate how well the alternatives are likely to meet the criteria and the weights that indicate the priorities of the criteria in achieving the decision goal. The priority weights and alternative scores are then combined to create summary scores that reflect the relative strengths and weaknesses of the alternatives relative to the goal of the decision. “What-if” sensitivity analyses can then be performed to explore the effects of changing the alternative scores and criteria priorities.

Ordinal methods—The simplest approach to value-based multi-criteria decision making is to create measurement scales by rank ordering the alternatives and the criteria. A number

of such methods are available. [7] They are called ordinal methods because the ranking process orders the decision elements being compared from most to least preferred.

One of the easiest to implement is a newly proposed method called the Multi-Attribute Global Inference of Quality (MAGIQ). [29] This method starts with a balance sheet of non-dominated alternatives. The criteria are then rank-ordered with regard to their priority relative to the decision goal and the alternatives are rank-ordered with regard to how well they meet each criterion. These rank orders are then converted into numeric scales using rank order centroids, a value that estimates the distance between adjacent ranks on a normalized scale running from 0 to 1. Rank order centroids are fairly easy to calculate but are easier to obtain by using a table of pre-calculated values. [30] Details about the calculation and a table of values are included in Appendix 2.

Ordinal method illustration—To illustrate the MAGIC technique, let's see how Mrs. Gray and her provider could use it to analyze her treatment decision. The procedure starts with the simplified balance sheet containing the non-dominated alternatives shown in Table II. Because the alternatives are already rank-ordered relative to the criteria, the only input required is the rank ordering of the criteria and sub-criteria. Mrs. Gray and her provider decide that Effectiveness is the highest priority, followed by Risk of Adverse Effects, and Cost. With regard to the sub-criteria, they judge halting disease progression more important than controlling symptoms and avoiding serious adverse effects more important than avoiding common side effects. The results of this process are shown in Table IV.

The next step is to convert the ranks into weights by substituting the appropriate rank order centroids. When there are three items the rank weights are 0.611, 0.2778, and 0.111. When there are two, the weights are 0.75 and 0.25. In the last step, final scores are calculated for each alternative by the weighted average method. To do this, the alternative's scores on the criteria are multiplied by the criteria priority weights and the results are summed. As shown in Table IV, the results indicate that, for Mrs. Gray, Drug A is the preferred treatment. Figure 2 illustrates the relative contributions of each decision criterion on the final alternative scores.

Comments—The basic premise of ordinal-based multi-criteria methods is that the process of rank ordering and weighting the criteria and alternatives should improve the soundness of the decision making process by assessing the decision makers' preferences about the pros and cons of the alternatives and the relative priorities of the decision criteria. This step should also improve communication among decision makers about important aspects of the choice being made. A major advantage of ordinal-based approaches is that the only input required of the decision makers is the rank ordering which is easy to understand and accomplish. It should also not be overly time consuming when added to a clinical consultation.

A major unresolved problem with ordinal ranking methods, however, is whether they adequately measure the decision makers' preferences. Ordinal forms of measurement indicate which items are better than others but provide no information about the magnitude of the differences. Other multi-criteria methods, such as those described below, use interval and ratio measurement scales that include this information and therefore more precisely quantify the decision maker's preferences. The accuracy with which full measurements scales can be derived from limited information is a fundamental question about rank order centroid weights and the approaches used to derive weights used other ordinal methods.

I am not aware of either any published medical applications or readily available computer software programs. These methods could, however, be easily implemented using a standard computer spreadsheet program.

Direct weighting—One of the most popular value-based methods is direct weighting. The simplest direct weighting method is to assign a number to every decision item being assessed. A typical approach is to assign a score of 1 to the lowest ranked item, and then determine scores for the others based on how much better or more important they are. These assessments also be made using a graphic interface. As with the ordinal methods, summary scores indicating how well the alternative meet the goal are typically created using the weighted average method. To ensure comparability and interpretation of results, all scores and priorities are frequently normalized by dividing each assigned score by sum of all scores so that they sum to one. The analysis can either be done in a top-down fashion, which starts with determining the criteria priorities, followed by assessments of the alternatives, or using a bottom-up approach that starts with the alternative assessments and ends with the criteria priorities.

Direct weighting example—Table V illustrates how a direct weighting approach could be applied to Mrs. Gray's treatment decision using a top-down approach. The clinical decision makers decide that Cost is the least important major criterion and give it a score of 1. They then judge Effectiveness 4 times more important, and Risk of Adverse Effects 3 times more important than Cost. The sum of these raw scores is 8. Dividing each individual score by their sum yields the following normalized priorities: effectiveness 0.50, risk of adverse effects 0.38, and cost 0.13.

This same process is then used to assign priorities to the sub-criteria under Effectiveness and Risk of Adverse Effects. They judge effectiveness in halting disease progression more important than symptom relief. Symptom relief is therefore assigned a raw score of 1. Halting disease progression is judged 4 times more important and assigned a raw score of 4. In terms of adverse effects, they judge the risk of serious adverse effects 5 times more important than common side effects. Normalized intra-criterion priorities for the sub-criteria in the same criterion family are then calculated by dividing the raw values by their sum. Global decision priority weights are then calculated by multiplying each localized intra-criterion priority by the priority weight of the parent criterion. Finally, scores for the alternatives relative to each criterion are calculated in the same fashion. All of these judgments and the resulting option scores are shown in Table V. Summary scores are then calculated using the weighted average method.

The end results of this process are also shown in Table V. Drug A is the preferred alternative. Figure 2 shows the relative contributions of each criterion toward the total scores.

Comment—The soundness of direct rating procedures is based on the premise that the process of assigning discrete scores to the decision options and criteria will make them more accurate measures of the decision maker's preferences. In addition, the process of quantifying criteria priorities and assessments of the options' strengths and weaknesses may facilitate better communication among the decision makers by making otherwise personal judgments explicit. Another advantage is that the quantification of these judgments makes it possible to easily examine the effects of changing criteria priorities and/or option rankings on the results. The major drawback with direct weighting is that it does not provide assurance that the assigned scores are valid and reliable indicators of decision maker's preferences and judgments.

Variations of the direct weighting method have been used in several patient decision aids and to investigate patient and clinicians' values for potential outcomes of rectal cancer treatment. [31–33] This method is well suited for implementation using a computer spreadsheet program. There are also several software programs available including *WEB-HIPRE*, a free web-based program available at <http://www.hipre.hut.fi/> and *Annalisa*, available at <http://www.cafeannalisa.org.uk/index.php>.

Multi-attribute utility analysis

Multiple criteria methods based on multi-attribute expected utility theory (MAUT), seek to improve the intensity of the decision making support beyond that provided by direct weighting methods by creating standardized scales called utility functions that measure how well the options meet the criteria and incorporating information about the variability of the options within each criterion in determining the criteria priorities. The most theoretically correct MAUT method is very difficult to implement in practice. For this reason, three simpler methods have been developed that very closely approximate the results of the original method: SMART (Simple Multi-Attribute Rating Technique), SMARTS (SMART using Swings), and SMARTER (SMART Exploiting Ranks).[30,34] All three methods take the same approach to structuring the decision problem, defining decision criteria, identifying options, generating utility functions, and creating summary scores. They differ in the method used to generate the criteria priority weights.

MAUT utility functions—A fundamental feature of MAUT is the generation and use of utility functions to evaluate how well the options meet the criteria. These functions differ from the simple normalized scales used in the direct weighting methods in that they transform the raw data from various criterion-specific measurement units to a common, dimensionless, interval-level scale running from 0 to 100. In its most basic form, the creation of a utility function involves three steps: a) determining the form of the utility function, b) defining the range of possible values the function should include, and c) mapping the options.

The simplest approach to creating these utility functions for information already characterized using a quantitative scale is to assume that they are linear. If so there are three basic scale types: 1) higher values are better than lower values, 2) lower values are better than higher ones, and 3) the best value falls somewhere between the two extremes. For types 1 and 2, all that is needed to define a linear utility function is to determine the minimum and maximum values of the range being included. A good general strategy is to use the full range of acceptable values for these endpoints, rather than just the range provided by the options that are being considered. In this way, the same function can be used repeatedly even if alternatives with different characteristics are included in the analysis. For type 3 scales, after the endpoints are chosen, it is also necessary to determine the location of the best value and whether the extremes are equally inferior or if one is better than the other.[30]

The use of a linear scale is appropriate as long as differences between similar intervals are judged to be equivalent for the entire range of possible values. For example, to determine if a linear function can be used to represent the effectiveness ratings of the medications in Mrs. Gray's decision, the value of a five point change would have to be judged the same for absolute differences in response rates between 0 and 5%, 50% and 55%, and 90% and 95%. As a general rule, even if these differences are not all judged equal, the use of a linear scale is acceptable as long as the ratio between any two intervals is less than or equal to 2:1. Ratios greater than this indicate that a non-linear scale will have to be created. [30] A full description of procedures used to create non-linear utility scales is beyond the scope of this paper; details can be found elsewhere. [34]

In some cases, decisions involve criteria that are not easily described using quantitative scales. In this case, each set of decision makers need to create an appropriate function. This is done using the direct weighting approach, as described above, to determine values for each alternative. The full utility function is then created in the standard fashion using these values in place of the pre-existing scale.

MAUT criteria priorities—As noted above, the MAUT methods differ in how the criteria priorities are derived. The initial method, SMART, used a direct weighting approach identical to the one described above except for the addition of a step that checks on the consistency of the judgments being made. This is done by sequentially eliminating the lowest ranked criterion and repeating the comparison procedure using the next lowest ranked criterion as the comparison standard. Judgments are consistent if they are either the same or very close for each set of comparisons. Inconsistent results suggest that decision makers have not made up their minds regarding the priorities of the criteria or a mistake has been made in performing the analysis. In this situation, more time should be taken until a more consistent set of priorities has been achieved. [34–35]

SMARTS and SMARTER were both created to improve the soundness of the original SMART procedure by ensuring that the derived criteria priorities incorporate the range of differences among the options. The underlying premise is that, for any particular decision, criteria that capture more of the differences among the options should play a greater role. Thus, a criterion where there is a 10-fold difference between the best and worst options should have a higher priority than one where the difference is only 2-fold. Of course, this approach to assigning priorities to decision criteria can be used for any value-based multi-criteria method. The difference is that this strategy is explicitly incorporated in these methods. The first modification using this procedure, SMARTS, has been replaced by the easier to use SMARTER procedure, so it won't be discussed further. [30]

SMARTER uses swing weights, a procedure similar to the one used to make the even swaps described earlier, to prioritize the criteria. The decision makers are asked to pretend they have an option that has the lowest possible score on all the criteria. They are then asked to indicate which criterion they would choose if they could improve the option's performance from worst to best on just one criterion. The chosen criterion is then given the highest priority. This process is then repeated using only the remaining criteria until all the criteria have been ranked. Priority weights are then assigned using the Rank Order Centroid Method described above in the Ordinal Methods section. [30]

When the decision model contains just one level of criteria, each criterion is prioritized directly. When the model includes both criteria and sub-criteria, the simplest way to proceed is to compare the proximate criteria, defined as the criteria and sub-criteria that are directly linked with the alternatives in the decision model. Parent criteria weights are determined by summing the weights of their related sub-criteria.

SMARTER method example—To illustrate the SMARTER method, I will show how Mrs. Gray and her provider could use it to help guide their decision making.

First, the raw data describing how well the alternatives meet the criteria have to be transformed into utility functions. In most cases this would be done beforehand, outside of the clinical encounter. To illustrate the process, let's assume that all functions meet the criteria for the linear approximation. The lower and upper bounds chosen for the decision criteria are: 40% and 100% for both effectiveness sub-criteria, 2% and 0 for serious adverse effects, 20% and 0 for common adverse effects, and \$150 and 0 for cost. Values for the decision alternatives are then determined by transforming the raw data onto these utility

scales. This is done using the following formula: $u = V \times (100/\Delta) - ((100/\Delta) \times \min)$, where u is transformed value, V equals the raw data value, Δ equals the difference between the maximum acceptable value and the minimum acceptable value, and \min equals the minimal acceptable value. The resulting utility functions are illustrated in Figure 3; the resulting option utilities are shown in Table VI.

The next step is to determine the criteria priorities. Because Mrs. Gray's decision involves both sub-criteria and criteria, the prioritization process is done by comparing the proximate criteria: the two sub-criteria under effectiveness, the two sub-criteria under adverse effects, and the criterion cost. Mrs. Gray and her provider choose effectiveness in preventing long term harm as the highest priority proximate criterion, followed by risk of serious side effects, monthly out of pocket cost, effectiveness in relieving symptoms, and risk of common side effects. This rank ordering is then used to assign rank order centroid weights. The option utilities and criteria priorities are then combined using the additive weighting method to generate summary scores for the alternatives.

As shown in Table VI, the results indicate that the differences among the three options are small; the scores for Drugs A, B₁, and C are 0.35, 0.34 and 0.31 respectively. Figure 2 shows the contributions of each criterion toward the summary scores. This result should prompt Mrs. Gray and her provider to explore the effects of changing their proximate criteria rankings and re-run the analysis after making any adjustments they think are warranted. If they decide that no adjustments are necessary, they should consider whether there is an additional consideration that was not included in the original analysis. An example would be the greater degree of uncertainty about the side effects of the newly released Drug C. If not, the best strategy would be to choose Drug A as the best option.

Comments—Multi-attribute utility analysis using the SMARTER method has several advantages for clinical decision support applications. It has a well-established theoretical background and, if pre-existing utility scales are used, the only input required of the clinical users is the swing-based ranking of the criteria. Along with these advantages, however, there are several potential drawbacks. The soundness of the SMARTER technique depends on decision makers' abilities to accurately make the swing judgments among the criteria, the extent to which linear value scales accurately represent their judgments, and the appropriateness of using criteria priority weights based on ordinal rankings. None of these assumptions has been tested in clinical decision support settings.

There has been at least one medical application of the SMARTER technique. [36] It is simple enough to be easily implemented using standard computer spreadsheet. A modification is also included in *WIN-HIPRE*.

The Analytic Hierarchy Process

Overview—Currently the most widely used multi-criteria method for both medical and non-medical applications is the Analytic Hierarchy Process (AHP). [8,11] The AHP was expressly created to be an easy to use decision support method capable of addressing a wide range of difficult decision problems including those that involve both “hard” data and less tangible considerations. Like the other methods in this category, it decomposes a decision problem, creates a quantitative comparison of the decision alternatives, and provides a format for performing “what-if” sensitivity analyses. Its level of decision support is comparable to that provided by MAUT. Differences from MAUT include the use of ratio level scales to rate the alternatives, the availability of several alternative methods for generating these scales, the use of a simple pairwise comparison procedure that is used throughout the analysis, and the routine calculation of a check on the internal consistency of these comparisons.

The AHP method—A classic AHP analysis starts with a balance sheet containing all non-dominated decision options. Like direct weighting, the scores regarding how well the options meet the criteria are determined directly from the data available. The difference is that, instead of assigning weights directly, they are determined by sequentially comparing all possible pairs of options in terms of their abilities to meet the criteria using a one to nine scale. Unlike MAUT, this same procedure can be applied equally well to both objective and more subjective criteria. After the comparisons are complete, they are combined using a matrix algebra calculation called the right principal eigenvector. This is equivalent to taking the average of both the direct comparisons that were made and the indirect comparisons they imply. The result is a normalized, ratio-level scale that reflects the judgments made between the alternatives. In addition to this scale, the pairwise comparison process also yields a measure of their consistency called the consistency ratio. A perfectly consistent set of judgments has a consistency ratio equal to zero. By convention, consistency ratios ≤ 0.10 are generally considered acceptable. In certain applied settings such as general population field studies ratios up to 0.15 are permissible. [37–39]

After the comparisons among the options are complete, the same pairwise comparison method is used to determine the priorities of the criteria relative to the decision goal and any sub-criteria relative to their parent criteria. When all the comparisons are complete, they are combined using the weighted additive method to create a normalized, summary scale that indicates how well the options have been judged to meet the decision goal.

Variations of this classic approach include a ratings method that uses the pairwise comparison process to assign weights to categories of alternatives - such as high, medium, or low - rather than individual options, two methods for calculating alternative scores depending on whether the goal of the analysis is to create a scale that reflects the pros and cons of the alternatives or to pick the best one, and the use of a network, rather than a hierarchical value tree, to conceptualize the decision. The ratings approach is especially useful when there are large numbers of alternatives under consideration. The different ways to assign alternative weights are used to avoid problems that can arise if two very similar alternatives are included in the analysis. The network approach, called the Analytic Network Process (ANP), is designed to avoid problems in assigning criteria priorities that could arise if the range of options available is not considered, the same issue that led to the SMARTS and SMARTER refinements of the original SMART technique. Further discussion of these AHP variations is beyond the scope of this paper, full descriptions are available elsewhere. [37,40–43]

AHP example—A classic AHP analysis of Mrs. Gray's decision problem would start after the balance sheet of non-dominated alternatives is created, as shown in Table II. The first step is to perform pairwise comparisons of the three options - Drug A to Drug B₁, Drug A to Drug C, and Drug B₁ to Drug C—relative the proximate decision criteria. Mrs. Gray's comparisons and the resulting alternative scores are illustrated in Table VII. In each case her consistency ratio is ≤ 0.10 , acceptable using the conventional cutoff for inconsistency.

After the alternative comparisons are complete, the same procedure is used to judge the importance of the sub-criteria relative to their parents and the major criteria in achieving the goal of the decision. The results indicate that Drug C is the best option with a score of 0.38, followed by Drug A (score 0.36), and Drug B₁ (score 0.25). These steps and the overall results of the analysis are also shown in Table VII. The relative contributions of the criteria toward the summary scores are illustrated in Figure 2.

Comment—The advantages of using the AHP to provide clinical decision support include its flexibility, ease of use, and strength of measurement. Its flexibility is due to the different

formats available. This makes it possible to adapt the same basic decision support method to different users and different circumstances. Its ease of use stems from the pairwise comparisons. People easily learn how to make these comparisons and the same method can be applied equally well to quantitative data and subjective considerations. Because the AHP is also a theory of measurement, the mathematical operations involved in the analysis are theoretically justified and assumption-free. [44] The routine use of the consistency ratio during the analysis helps to help users avoid making technical mistakes and monitor the quality of the analysis. The main drawback of using the AHP for clinical decision support is that the pairwise comparison process is time consuming which can make it challenging to implement in clinical settings.

A recent review of medical applications of the AHP identified 50 published articles. [11] The calculations required to perform an AHP analysis can be easily programmed into a computer spreadsheet. *WEB-HIPRE* supports AHP analyses; a number of commercial and non-commercial software programs are also available including a free program called *Superdecisions* that also supports ANP analyses that is available at www.superdecisions.com.

Conclusion

Full realization of the goal to provide patient-centered, evidence-based care throughout the healthcare system will require effective and efficient ways for enabling patients and providers to routinely carry out the necessary tasks. Multi-criteria decision making techniques are uniquely well-suited for this purpose, suggesting that they could serve as the foundation for a new generation of clinical decision support systems designed to support consistent, high quality medical care.

All of the methods described here are capable of helping patients and clinicians like Mrs. Gray and her provider gain a deeper understanding of clinical decisions they face and reduce their initial state of uncertainty about the best course of action. The key to making these methods useful in clinical practice is making sure they both meet the needs of the intended users and provide substantive decision making support.

As noted, there are unanswered questions about the clinical applicability of each method that are primarily due to an unavoidable inverse relationship between ease of use and the amount of decision support they provide. It is likely that the usefulness of every method will be affected by the nature of the clinical problem being addressed and individual patient needs and capabilities. More difficult or important decisions are better suited for more intensive deliberation and decision support. A technically high level of decision support that does not make sense to the decision maker, however, is of no benefit. Whether or not multi-criteria decision making will ultimately serve as the cornerstone of the next generation of decision support interventions is currently unknown as research into the actual usefulness and feasibility of multi-criteria based clinical decision support systems is in its infancy. However, the potential exists. Additional research to determine the role of multi-criteria methods in healthcare decision making is warranted.

Acknowledgments

This work was supported by grant 5K24HL093488-02 from the National Heart Lung and Blood Institute, US National Institutes of Health.

References

1. Committee on quality of health care in America IoM. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, D.C.: National Academy Press; 2001.
2. National Healthcare Quality Report. 2008. [2010 August 11]; Available from: <http://www.ahrq.gov/qual/nhqr08/Key.htm>
3. Elwyn G, Edwards A, Gwyn R, Grol R. Towards a feasible model for shared decision making: focus group study with general practice registrars. *BMJ* 1999;319:753–756. [PubMed: 10488002]
4. Bodenheimer T, Wagner EH, Grumbach K. Improving Primary Care for Patients With Chronic Illness: The Chronic Care Model, Part 2. *JAMA* 2002;288(15):1909–1914. [PubMed: 12377092]
5. O'Connor AM, Wennberg JE, Legare F, Llewellyn-Thomas HA, Moulton BW, Sepucha KR, et al. Toward The 'Tipping Point': Decision Aids And Informed Patient Choice. *Health Aff* 2007;26(3): 716–725.
6. Belton, V.; Stewart, TJ. Multiple Criteria Decision Analysis. Boston/Dordrecht/London: Kluwer Academic Publishers; 2002.
7. Figueira, J.; Greco, S.; Ehrgott, M. State of the art surveys. New York: Springer; 2005. Multiple Criteria Decision Analysis.
8. Wallenius J, Dyer JS, Fishburn PC, Steuer RE, Zionts S, Deb K. Multiple Criteria Decision Making, Multiattribute Utility Theory: Recent Accomplishments and What Lies Ahead. *Management Science* 2008;54(7):1336–1349.
9. Baltussen R, Niessen L. Priority setting of health interventions: the need for multi-criteria decision analysis. *Cost Effectiveness and Resource Allocation* 2006;4(1):14–14. [PubMed: 16923181]
10. Goetghebeur MM, Wagner M, Khoury H, Levitt RJ, Erickson LJ, Rindress D. Evidence and Value: Impact on DEcisionMaking--the EVIDEM framework and potential applications. *BMC Health Serv Res* 2008;8:270. [PubMed: 19102752]
11. Liberatore MJ, Nydick RL. The analytic hierarchy process in medical and health care decision making: A literature review. *European Journal of Operational Research* 2008;189(1):194–207.
12. Sloane EB, Liberatore MJ, Nydick RL. Medical decision support using the Analytic Hierarchy Process. *Journal of Healthcare Information Management: JHIM* 2002;16(4):38–43. [PubMed: 12365298]
13. Sloane EB, Liberatore MJ, Nydick RL, Luo W, Chung QB. Using the analytic hierarchy process as a clinical engineering tool to facilitate an iterative, multidisciplinary, microeconomic health technology assessment. *Computers & Operations Research* 2003;30(10):1447–1465.
14. Pieterse AH, Berkers F, Baas-Thijssen MC, Marijnen CA, Stiggelbout AM. Adaptive Conjoint Analysis as individual preference assessment tool: feasibility through the internet and reliability of preferences. *Patient Education & Counseling* 2010;78(2):224–233. [PubMed: 19581069]
15. Mele NL. Conjoint Analysis. *Nursing Research* 2008;57(3):220–224. [PubMed: 18496108]
16. Ahmed SF, Smith WA, Blamires C. Facilitating and understanding the family's choice of injection device for growth hormone therapy by using conjoint analysis. *Archives of disease in childhood* 2008;93(2):110–110. [PubMed: 17412745]
17. Fisher K, Orkin F, Frazer C. Utilizing conjoint analysis to explicate health care decision making by emergency department nurses: a feasibility study. *Applied Nursing Research* 2010;23(1):30–35. [PubMed: 20122508]
18. Ryan M, Farrar S. Using conjoint analysis to elicit preferences for health care. *BMJ* 2000 Jun 3;320(7248):1530–1533. [PubMed: 10834905]
19. Breslin M, Mullan RJ, Montori VM. The design of a decision aid about diabetes medications for use during the consultation with patients with type 2 diabetes. *Patient Education and Counseling* 2008;73(3):465–472. [PubMed: 18771876]
20. Few, S. Information dashboard design. O'Reilly; 2005.
21. Mullan RJ, Montori VM, Shah ND, Christianson TJH, Bryant SC, Guyatt GH, et al. The diabetes mellitus medication choice decision aid: a randomized trial. *Archives of Internal Medicine* 2009;169(17):1560–1568. [PubMed: 19786674]
22. MarewskiJNGaissmaierWGigerenzerGGood judgments do not require complex cognition. *Cogn Process* 2010 May 11;2103121 [PubMed: 19784854]

23. Eddy DM. Comparing benefits and harms: the balance sheet. *JAMA: The Journal of the American Medical Association* 1990;263(18):2493–2505. [PubMed: 2329639]
24. Dougherty MR, Franco-Watkins AM, Thomas R. Psychological plausibility of the theory of probabilistic mental models and the fast and frugal heuristics. *Psychological Review* 2008;115(1): 199–213. [PubMed: 18211192]
25. Hastie, R.; Dawes, RM. *Rational choice in an uncertain world*. Thousand Oaks, CA: Sage Publications, Inc; 2001.
26. Hammond JS, Keeney RL, Raiffa H. Even swaps: a rational method for making trade-offs. *Harvard Business Review* 1998;76(2):137–138, 143–148, 150–137–138, 143–148, 150. [PubMed: 10177863]
27. Hammond, JS.; Keeney, RL.; Raiffa, H. *Smart Choices: A Practical Guide to Making Better Decisions*. Boston: Harvard Business School Press; 1999.
28. Mustajoki J, Hamalainen RP. Smart-Swaps - A decision support system for multicriteria decision analysis with the even swaps method. *Decision Support Systems* 2007;44:313–325.
29. McCaffrey JD. Using the Multi-Attribute Global Inference of Quality (MAGIQ). *Technique for Software Testing* 2009;2009:738–742.
30. Edwards W, Barron FH. SMARTS and SMARTER: Improved Simple Methods for Multiattribute Utility Measurement. *Organizational Behavior and Human Decision Processes* 1994;60(3):306–325.
31. O'Connor AM, Tugwell P, Wells GA, Elmslie T, Jolly E, Hollingworth G, et al. A decision aid for women considering hormone therapy after menopause: decision support framework and evaluation. *Patient Education and Counseling* 1998;33(3):267–279. [PubMed: 9731164]
32. Man-Son-Hing M, Laupacis A, O'Connor AM, Hart RG, Feldman G, Blackshear JL, et al. Development of a decision aid for patients with atrial fibrillation who are considering antithrombotic therapy. *Journal of General Internal Medicine* 2000;15(10):723–730. [PubMed: 11089716]
33. Masya LM, Young JM, Solomon MJ, Harrison JD, Dennis RJ, Salkeld GP. Preferences for outcomes of treatment for rectal cancer: patient and clinician utilities and their application in an interactive computer-based decision aid. *Diseases of the Colon and Rectum* 2009;52(12):1994–2002. [PubMed: 19934920]
34. Von Winterfeldt, D.; Edwards, W. *Decision analysis and behavioral research*. Cambridge: Cambridge University Press; 1986.
35. Yoon, KP.; Hwang, DC-L. *Multiple Attribute Decision Making: An Introduction*. Sage Publications, Inc; 1995.
36. de Bock GH, Reijneveld SA, van Houwelingen JC, Knottnerus JA, Kievit J. Multiattribute utility scores for predicting family physicians' decisions regarding sinusitis. *Med Decis Making* 1999 jan-mar;19:58–65. [PubMed: 9917021]
37. Forman EH, Gass SI. The Analytic Hierarchy Process - An Exposition. *Operations Research* 2001;49:469–486.
38. Katsumura Y, Yasunaga H, Imamura T, Ohe K, Oyama H. Relationship between risk information on total colonoscopy and patient preferences for colorectal cancer screening options: Analysis using the Analytic Hierarchy Process. *BMC Health Services Research* 2008;8(1):106–106. [PubMed: 18492286]
39. Sato J. Comparison between multiple-choice and analytic hierarchy process: measuring human perception. *International Transactions in Operational Research* 2004;11(1):77–86.
40. Dolan JG, Isselhardt BJ, Cappuccio JD. The analytic hierarchy process in medical decision making: a tutorial. *Medical Decision Making: An International Journal of the Society for Medical Decision Making* 1989;9(1):40–50. [PubMed: 2643019]
41. Saaty TL. How to make a decision: The analytic hierarchy process. *European Journal of Operational Research* 1990;48(1):9–26.
42. Saaty TL. How to make a decision: The Analytic Hierarchy Process. *Interfaces* 1994;24:19–43.
43. Saaty, TL. *Decision Making for Leaders*. Pittsburgh, PA: RWS Publications; 2001.
44. Gass SI. Model world: When is a number a number? *Interfaces* 2001;31(5):93–103.

Appendix 1: Glossary of multi-criteria terms

Additive weighting method

A method of determining the overall scores of decision options analogous to calculating a weighted average. Calculated by multiplying the options' scores on the criteria times the weights of the criteria and summing across all criteria.

Alternative

A course of action being actively considered as part of a decision making process. Used interchangeably with option.

Attribute

A characteristic or feature of a decision option.

Common side effect

An adverse effect of a medication that causes temporary symptoms such as rash or headache that generally require minimal treatment for control and do not necessarily indicate that the causative medication needs to be discontinued.

Compensatory method

Decision making methods that incorporate information from all the decision criteria.

Consistency

The extent to which a set of related judgments are internally consistent with each other.

Criterion

Consideration being used to select a preferred alternative in making a decision. Usually refers to attributes or characteristics of the alternatives that determine their desirability. Used interchangeably with objective.

Decision elements

The components of a decision including the decision goal, the criteria and sub-criteria, and the alternatives being considered.

Decision model

A graphic representation of a decision that lists the options being considered and the considerations being used to compare the options. Sometimes a decision goal is also included. The format used is variable. Common arrangements include hierarchies and networks.

Decomposed approach

A method of decision analysis that breaks a decision down into separate elements such as alternatives and criteria.

Dimension

An important consideration in making a decision.

Dominated alternative

A decision alternative that has no unique advantages compared to the other alternatives being considered and is always inferior to at least one other alternative for every dimension.

Interval scale

A scale of measurement where there is a defined distance between any two points on the scale.

Non-compensatory method

Decision making methods that do not incorporate information from all the decision criteria.

Objective

Consideration being used to select a preferred alternative in making a decision. Usually refers to attributes or characteristics of the alternatives that determine their desirability. Used interchangeably with criterion.

Option

A course of action being actively considered as part of a decision making process. Used interchangeably with alternative.

Ordinal scale

A scale of measurement that involves assigning items to higher or lower ranks. No information is provided about the magnitude of the differences between the ranks.

Pairwise

A procedure involving a pair of items. Most commonly used to describe a process of making judgments between two decision elements.

Proximate criteria

Criteria and sub-criteria that are directly linked to decision options in a decision model.

Ratio level scale

A scale of measurement where there is both a defined distance between any two points on the scale and an absolute zero point. This is the only type of scale where ratios between numbers have meaning.

Score

A number representing the relative importance of a decision element in making a choice between a set of options. Most commonly applied to options.

Serious side effect

An adverse effect of a medication that requires substantial treatment for control, could result in permanent effects, and requires discontinuation of the causative drug.

Value Tree

A hierarchical decision model with the goal of the decision on the highest level, the options on the lowest level, and the criteria and sub-criteria being used to compare the options relative to the goal in the middle.

Value-based methods

Multi-criteria methods that develop quantitative measures of how well the options fulfill the criteria and the relative impacts of the criteria in achieving the goal of the decision.

Weight

A number representing the relative importance of a decision element in making a choice between a set of options. Most commonly applied to criteria and sub-criteria.

Appendix 2. Rank-order Centroids [30]

Calculation

Assume that there are K items that need to be weighted and that they are rank-ordered so that $w_1 > w_2 > w_3 > w_K$. The rank order centroid of any item i is equal to:

$$w_k = (1/K) \sum_{i=k}^K (1/i)$$

For example, if there are three items to be weighted, $K = 3$. The rank order centroid of the items are then calculated using the following formulas:

$$w_1 = (1 + 1/2 + 1/3) / 3 = 0.61$$

$$w_2 = (0 + 1/2 + 1/3) / 3 = 0.28$$

$$w_3 = (0 + 0 + 1/3) / 3 = 0.11$$

When there are ties, the average of the weights for the tied places is used. For example, if there is a tie for the second best item, the mean of the weights for ranks 2 and 3 is used for both of the tied items.

Table of rank order centroids

Rank	Number of items							
	2	3	4	5	6	7	8	9
1	0.75	0.61	0.52	0.46	0.41	0.37	0.34	0.31
2	0.25	0.28	0.27	0.26	0.24	0.23	0.21	0.20
3		0.11	0.15	0.16	0.16	0.16	0.15	0.15
4			0.06	0.09	0.10	0.11	0.11	0.11
5				0.04	0.06	0.07	0.08	0.08
6					0.03	0.04	0.05	0.06
7						0.02	0.03	0.04
8							0.02	0.03
9								0.01

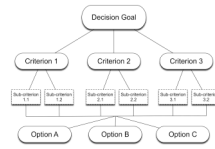


Figure 1. A generic multi-criteria decision-making model consisting of a goal, criteria, sub-criteria, and several options.

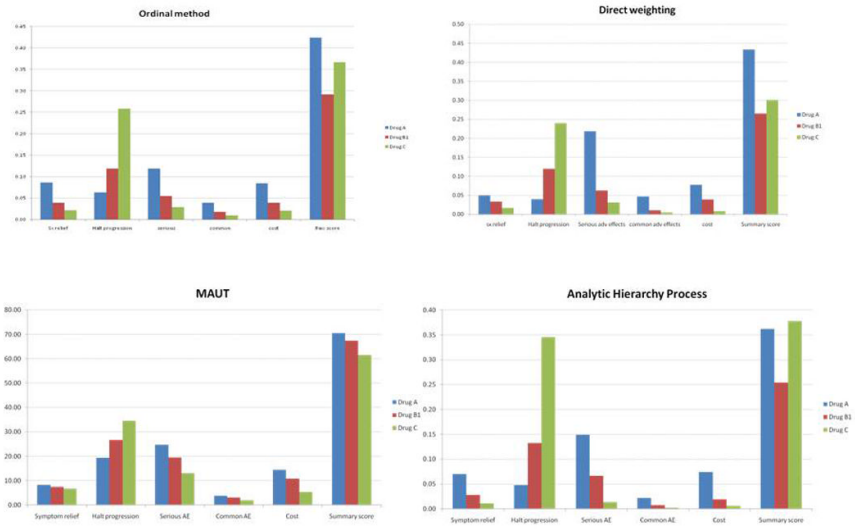


Figure 2. The results of Mrs. Gray's multi-criteria analysis using the four value-based methods. The summary scores indicate the overall results. The criteria-specific scores indicate the contribution of each criterion toward the summary score based on the priority weights and the criterion-specific performance of each alternative.

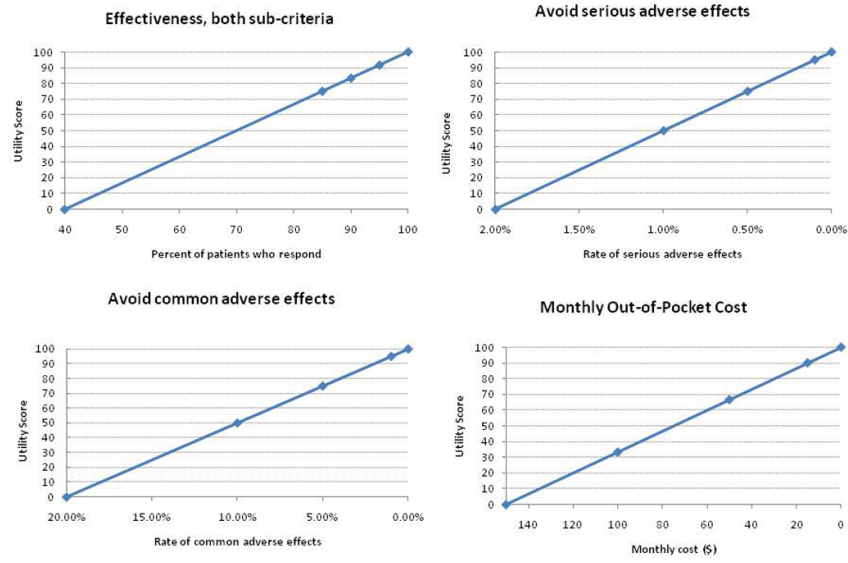


Figure 3. The utility scales generated from the data about the alternative drugs contained in the balance sheet.

Table 1

A balance sheet

Drug	Effectiveness		Risk of Adverse Effects				Cost			
	Symptom relief		Halt disease progression		Reported rate of serious side effects		Reported rate of common side effects		Monthly out-of-pocket cost	
	Response rate	Ranking	Response rate	Ranking	Rate	Ranking	Rate	Ranking	Cost	Ranking
A	95%	1	65%	4	0.1%	1	1%	1	\$15	1
B ₁	90%	2	75%	2	0.5%	2	5%	2	\$50	2
B ₂ ^a	80%	4	75%	2	0.5%	2	8%	3	\$75	3
C	85%	3	85%	1	1%	4	10%	4	\$100	4

^aDrug B₂ is a dominated treatment choice because it is never the best option and always equal to or inferior to Drug B₁. Unless other criteria are important in choosing among these treatment alternatives, it should be eliminated from further consideration.

Table II

Simplified balance sheet after removal of the dominated option

Drug	Effectiveness		Risk of Adverse Effects		Cost	
	Symptom relief	Halt disease progression	Reported rate of serious side effects	Reported rate of common side effects	Monthly out-of-pocket cost	Ranking
A	95%	65%	0.1%	1%	\$15	1
B ₁	90%	75%	0.5%	5%	\$50	2
C	85%	85%	1%	10%	\$100	3

Table III

Revised balance sheet including swaps.

Drug	Effectiveness			Risk of Adverse Effects			Cost
	Symptom relief	Halt disease progression	Reported rate of serious side effects	Reported rate of common side effects	Reported rate of serious side effects	Monthly out-of-pocket cost	
Swap 1: Increase cost of drug A from \$15 to \$50 in return for increase in halting disease progression effectiveness from 65% to 75%. Drug A now dominates Drug B ₁ which can now be eliminated from further consideration.							
A	95%	75%	0.1%	1%	1%	\$50	1
B ₁	90%	75%	0.5%	5%	2%	\$50	1
C	85%	85%	1%	10%	3%	\$100	2
Swap 2: Decrease Drug symptom relief effectiveness from 95% to 85% in return for increase in halting disease progression effectiveness from 75% to 80%.							
A	85%	80%	0.1%	1%	1%	\$50	1
C	85%	85%	1%	10%	2%	\$100	2
Swap 3: Increase Drug A common side effect rate from 1% to 5% and cost from \$50 to \$75 in return for increase in halting disease progression effectiveness from 80% to 85%. Drug A now dominates Drug C.							
A	85%	85%	0.1%	5%	1%	\$75	1
C	85%	85%	1%	10%	2%	\$100	2

Table IV

Multi-criteria analysis using the ordinal MAGIQ method.

	Effectiveness		Risk of Adverse Effects.		Cost	
	Symptom relief	Halt disease progression	Serious	Common		
Intra-criterion priority	#2, weight 0.25	#1, weight 0.75	#1, weight 0.75	#2, weight 0.25	Priority #3, weight 0.15	
Global decision priority	$0.61 \times 0.25 = 0.15$	$0.61 \times 0.75 = 0.46$	$0.28 \times 0.75 = 0.21$	$0.28 \times 0.25 = 0.07$	$0.15 \times 1 = 0.15$	Summary Scores ^a
	Alternative ranks and scores					Raw
Drug A	#1 - 0.61	#3 - 0.15	#1 - 0.61	#1 - 0.61	#1 - 0.61	0.42
Drug B ₁	#2 - 0.22	#2 - 0.28	#2 - 0.28	#2 - 0.28	#2 - 0.28	0.29
Drug C	#3 - 0.15	#1 - 0.61	#3 - 0.15	#3 - 0.15	#3 - 0.15	0.37
						Normalized
						0.39
						0.27
						0.34

^aSummary score calculated by multiplying alternative scores for each criterion times the criterion priorities and then normalizing by dividing each score the sum of scores. For example, the raw summary score of Drug A = $(0.15 \times 0.61) + (0.46 \times 0.15) + (0.21 \times 0.61) + (0.07 \times 0.61) + (0.15 \times 0.61)$; the normalized score = $0.42 / (0.42 + 0.29 + 0.37)$.

Table V

Direct weighting example

Calculation of criteria and sub-criteria priorities					
Criterion	Effectiveness		Adverse Effects		Cost
Priority judgment ^a	4		3		1
Priority weight ^b	0.5		0.38		0.13
Sub-criterion	Symptom relief	Halt disease progression	Serious	Common	
Intra-criterion priority judgment ^c	1	4	5	1	
Intra-criterion priority weight	0.2	0.8	0.83	0.17	1
Global priority ^d	0.10		0.31	0.06	0.13 ^e

Direct weighting of the options relative to the criteria										
Proximate criterion	Symptom relief		Halt disease progression		Serious Adverse effects		Common Adverse effects		Cost	
	Judgments ^a	Score ^b	Judgments	Score	Judgments	Score	Judgments	Score	Judgments	Score
Drug A	3	0.50	1	0.1	7	0.7	9	0.75	10	0.63
Drug B ₁	2	0.33	3	0.3	2	0.2	2	0.17	5	0.31
Drug C	1	0.17	6	0.6	1	0.1	1	0.08	1	0.06

Criterion	Halt disease progression		Serious adverse effects		Common adverse effects		Cost
	Symptom relief	Score	Judgments	Score	Judgments	Score	
Criterion priority	0.1	0.4	0.31	0.06	0.13		
Option Scores							
Drug A	0.05	0.04	0.22	0.05	0.08	0.43	
Drug B ₁	0.03	0.12	0.06	0.01	0.04	0.27	
Drug C	0.02	0.24	0.03	0.01	0.01	0.30	
							Summary Score ^a

^aPriority judgments refer to the relative importance of each criterion in making the decision. The least preferred option has a score of one. The other option scores indicate how much more preferred they are than the least preferred option.

^bPriority weights are calculated by dividing each criterion's priority judgment by the sum of all judgments.

^c Intra-criterion priority judgments refer to the relative importance of sub-criteria with respect to their parent criterion.

^d Global priority refers to the priority of the sub-criteria relative to the goal. A global priority is calculated by multiplying the intra-criterion priority weight times the priority weight of the parent criterion.

^e Note that the cost criterion has no sub-criteria. The full priority of the parent criterion is therefore used to weight the options.

^a Judgments refer to the relative preferences of each option for the indicated criterion. The least preferred option has a score of 1. The other option score indicate how much more preferred they are than the least preferred option.

^b Scores are calculated by dividing each option's judgment value by the sum of all judgment values for each criterion.

^a Raw summary score is determined by multiplying each drug's score on a priority times the criterion priority and summing across all criteria. For example, Drug A summary score = $(0.50 \times 0.1) + (0.10 \times .4) + (0.70 \times 0.31) + (0.75 \times 0.06) + (0.63 \times 0.13)$.

Table VI

Illustration of SMARTER MAUT analysis

Major criterion	Effectiveness		Adverse Effects				Cost		Normalized Total score ^a
	Symptom relief	Halt disease progression	Serious	Common	Original data	Utility score	Original data	Utility score	
Priority	0.55		0.3				0.16		
Priority rank	4	1	2	5			3		
Priority	0.09	0.46	0.26	0.04	0.16				
Drug A	95%	92	0.1%	95	1%	95	\$ 15	90	0.35
Drug B1	85%	83	0.5%	75	5%	75	\$ 50	67	0.34
Drug C	85%	75	1.0%	50	10%	50	\$ 100	33	0.31

^aTotal score calculated by multiplying each option's rank order centroid times the utility score for each criterion, summing over all criteria, and then normalizing the results. For example, Drug A score (before normalization) = $(0.09 \times 92) + (0.46 \times 42) + (0.26 \times 95) + (0.04 \times 95) + (0.16 \times 90)$

Table VII

Analytic Hierarchy Process (AHP) example.

Effectiveness				Adverse effects						Cost			
Symptom relief		Halt disease progression		Serious		Common		Cost					
Pref	SOP ^a	Score	Pref	SOP	Score	Pref	SOP	Score	Pref	SOP	Score		
A > B ₁	3	A = 0.64	B ₁ > A	3	A = 0.09	A > B ₁	3	A = 0.65	A > B ₁	4	A = 0.70		
A > C	5	B ₁ = 0.26	C > A	5	B ₁ = 0.25	A > C	8	B ₁ = 0.29	A > C	8	B ₁ = 0.24		
B ₁ = C	3	C = 0.10	C > B ₁	3	C = 0.65	B ₁ = C	6	C = 0.06	B ₁ = C	5	C = 0.06		
Consistency ratio 0.04			Consistency ratio 0.02			Consistency ratio 0.07			Consistency ratio 0.09			Consistency ratio 0.07	

Sub-criteria comparisons			
Effectiveness		Adverse effects	
Preference	SOP	Priority Scores	Priority Scores
Halt disease progression > Symptom control	5	Halt progression = 0.83 Symptom control = 0.17 Consistency ratio = 0	Serious > Common 7 Serious = 0.88 Common = 0.12 Consistency ratio = 0
Criteria comparisons			
Preference	SOP		Priority scores
Effectiveness > Adverse Effects	3		Effectiveness = 0.64 Adverse Effects = 0.26 Cost = 0.10
Effectiveness > Cost	5		
Adverse Effects > Cost	3		
Consistency ratio = 0.03			

Synthesis			
Major Criterion	Effectiveness	Adverse Effects	Cost
Priority Score	0.64	0.26	0.1
Intra-criterion priority	Symptom relief	Halt disease progression	Common
	0.17	0.83	0.88
		0.12	1

Global priority	0.11	0.53	0.23	0.03	0.1	Total Score ^a
	Option Scores					
Drug A	0.64	0.09	0.65	0.70	0.74	0.36
Drug B ₁	0.26	0.25	0.29	0.24	0.19	0.25
Drug C	0.10	0.65	0.06	0.06	0.06	0.38

Abbreviations: Pref = preference, SOP = strength of preference, > = more preferred, < = less preferred

^a Strength of preference scores are made on a nine point scale ranging from equal to extremely preferred or more important.

^a The total score is calculated by multiplying each option's score times the global priority for each criterion, summing over all criteria, and then normalizing the results. For example, Drug A's score = $(0.64 \times 0.11) + (0.09 \times 0.53) + (0.65 \times 0.23) + (0.70 \times 0.03) + (0.74 \times 0.1)$.