Original Article

MDM Medical Decision Making

Medical Decision Making
1–10

DOI: 10.1177/0272989X20915241
journals.sagepub.com/home/mdm
SAGE

Framework for Drug Formulary Decision Using Multiple-Criteria Decision Analysis

Vusal Babashov, Sarah Ben Amor, and Gilles Reinhardt

Background. Reviewing drugs to determine coverage or reimbursement level is a complex process that involves significant time and expertise. Review boards gather evidence from the submission provided, input from clinicians and patients, and results of clinical and economic reviews. This information consists of assessments on multiple criteria that often conflict with one another. Multiple-criteria decision analysis (MCDA) includes methods to address complex decision making problems with conflicting objectives and criteria. We propose an MCDA approach that infers a utility model based on reviews of previously submitted drugs. **Methods.** We use a recent extension of the UTilitiés Additives DIScriminantes approach, UTADIS^{GMS}. This disaggregation approach deconstructs a portfolio of elements such as a set of drugs that have been reviewed and for which a decision has been made. It derives global and marginal utility functions that are consistent with the preferences exhibited by the review boards in their recommendations. We apply the method to oncology drugs reviewed in Canada between 2011 and 2017. We also illustrate how to conduct scenario analyses and predict the coverage decisions for new drugs. **Results.** Applying the method yields a utility value for each submission along with a set of thresholds that partition the utility values based on the submission outcomes. Scenario analyses illustrate the predictive ability of the method. **Conclusion.** Preference disaggregation is an indirect way of eliciting an additive global utility value function. It requires less of a cognitive effort from the decision making bodies because it infers preferences from the data rather than relying on direct assessments of model parameters. We illustrate how it can be applied to validate existing decisions and to predict the recommendation of a new drug.

Keywords
decision support, formulary design, multiple-criteria decision analysis

Date received: March 27, 2019; accepted: February 19, 2020

A formulary is a list of drugs that public or private insurance plans approve for use by their beneficiaries. Requests to add a new drug to a formulary may come from manufacturers, patient groups, or clinicians.[1] The submitter forwards a file to a specialized review board or a pharmacy and therapeutics committee. The board or committee ultimately issues an inclusion or a reimbursement recommendation (i.e., full approval, conditional approval, rejection) based on the submitted file, on additional input from clinicians and patients, and on results from clinical and economic reviews.

We focus on Canadian oncology drug recommendations. Provincial cancer centers in Canada manage their oncology drug formularies based on the recommendations from the pan-Canadian Oncology Drug Review (pCODR). New submissions undergo the pCODR review process and obtain a reimbursement recommendation before provincial plans and cancer agencies consider funding

Telfer School of Management, University of Ottawa, Ottawa, ON, Canada (VB, SBA, GR). The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The author(s) received no financial support for the research, authorship, and/or publication of this article.

Corresponding Author
Gilles Reinhardt, Telfer School of Management, University of Ottawa, 55 Laurier Ave E, Ottawa, ON K1N 6N5, Canada (gilles.reinhardt@telfer.uottawa.ca).

them. Decision reports are publicly available on the pCODR website.[2]

After accepting to review a submission, the board sets up an expert review team that evaluates the drug along 4 dimensions: clinical benefit, economic efficiency, patient-based values, and feasibility.[1] It then aggregates the information and formulates an initial recommendation, which becomes final subsequent to feedback from the stakeholders involved in the process.

This complex process takes months to complete and involves dozens of panelists, clinicians, scientists, and community members. They form a large pool of diverse expertise, and their input to the review board includes measurements, appraisals, and arguments derived from a multiplicity of sources. Aggregating and transforming that input into a recommendation is also challenging, since it involves debating over several independent and conflicting factors before reaching a consensus.

Systematic decision support methods can facilitate this process. For instance, several optimization-based approaches model listing recommendation decisions. Stinnett and Paltiel[3] use a framework that maximizes effectiveness of a formulary subject to budgetary and ethical constraints. Their model addresses some of the shortcomings of the standard cost-effectiveness analysis.[4] Olmstead and Zeckhauser,[5] inspired by the menu-setting approach, couch the formulary problem as a model that maximizes total consumer welfare or social surplus, subject to constraints on the total pharmaceutical budget over different patient groups and medical conditions. Truong[6] extends this approach by incorporating patient choice, stochastic utility, and multiple drug categories and shows that selecting a drug based on the incremental cost-effectiveness ratio (ICER) generally results in suboptimal formularies. However, concrete formulary listing decisions require thorough analysis on many criteria, which limits the applicability of optimization-based models.

Multiple-criteria decision analysis (MCDA) is a field of operations research that addresses decision making problems with conflicting objectives and criteria. It provides recommendations according to one of the following problematics: choice, ranking, or sorting.[7] The sorting problematic concerns the assignment of the alternatives to predefined ordered classes, which is the goal of our study.

MCDA has gained traction in assessing health technologies and interventions.[8] A recent task force report from the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) on emerging good practices in MCDA outlines different health care settings in which MCDA tools can be applied and provides guidelines on implementation.[9,10] As a result, health technology assessment (HTA) agencies around the world have started using MCDA.[11,12]

MCDA models traditionally used a direct approach in which criteria parameters, or *preferences*, are given. Those models include the value system approach (in which *alternatives* are scored based on their criteria assessments), the outranking approach (based on pairwise comparisons of alternatives), and the multiobjective optimization approach (suitable when the set of alternatives is continuous rather than discrete).[13] The disaggregation paradigm is an indirect approach that disentangles previous decisions. Unlike its direct counterparts, that paradigm infers a model that captures the preferences of the decision making bodies.[14] In so doing, it circumvents the challenges of scaling and weighting criteria that HTA agencies have identified as a key concern with the implementation and interpretation of MCDA.[15,16]

A disaggregation approach is similar to machine learning since both are data driven. Although they produce comparable results in many instances, there are important differences in terms of dimensionality, inconsistencies, and validation.[17,18] Regarding dimensionality, machine learning operates with large data sets, whereas disaggregation approaches have sets with only a limited number of alternatives.[19] In our pCODR context, we note that committees consider a rather limited number of previous decisions, in addition to economic and clinical evidence, when issuing recommendations. We address the differences with respect to inconsistencies and validation in the Methods and Results sections.

Our goal in this article is to tailor a specific MCDA tool of the disaggregation paradigm, UTADIS^{GMS} (UTilités Additives DIScriminantes, with the superscript referring to its authors),[20] and illustrate with a set of existing recommendations how it can positively contribute to the pCODR review process. UTADIS^{GMS} takes as input a set of submissions that have already been decided upon and criteria assessments, which may be qualitative or quantitative. It then infers parameters that model the contribution of each criterion considered in the analysis and produces a global utility value for the submissions that have been decided upon that is compatible with their respective recommendation.

Since recommendations are known a priori, the model analyzes how the criteria contribute to the recommendations and estimates a representative global utility function. Applying the function to the set of submissions

yields the global utility value for each. The method also determines thresholds that partition the range of the utility values into ordered classes that cluster recommendation categories for each alternative, for instance, "full approval," "conditional approval," and "rejection."

This method provides valuable decision making support. It ensures consistency of outcomes by implicitly inferring the contribution of each criterion to the recommendation and, in so doing, validates existing decisions. Moreover, we can leverage its predictive ability for alternatives without decisions with vectors of criteria assessments that model those potential new drugs.

We structure the remainder of this article as follows. In the Methods section, we detail the UTADIS methodology and its extension, UTADIS^{GMS}. We also describe the data set we use in our analysis of Canadian oncology drug reviews. In the Results section, we demonstrate the predictive capability of the model by developing 3 "scenarios" (potential new drugs) and analyzing the outcomes. We highlight our contributions, discuss the limitations of our model and findings, and point out avenues for future research in the Discussion section.

Methods

UTADIS was first presented by Devaud et al.[21] to develop a criteria aggregation model to determine the classification of alternatives. It is a version of the well-established UTilités Additives method (UTA) used for ranking a set of alternatives from best to worst. An additive model is deemed sufficient to capture the decision maker's preferences when the criteria are independent.[22] UTADIS has been applied in financial decision making,[23,24] project and portfolio selection,[25] and business performance evaluation,[26] notably in health care organizations.[27] Some extensions have also been developed.[28–30]

Greco et al.[20] extended the original UTADIS method and built a multicriteria sorting model based on $\mathcal{U}_{A^R}$, the set of all global additive value functions. Those are composed of monotonic and possibly nonlinear marginal value functions that are compatible with the provided preference information from the decision making bodies. This extension, UTADIS^{GMS}, is a classification method based on ordinal regression. It considers value functions that can be convex or concave rather than piecewise linear. This so-called epsilon maximization approach, which we describe below, ensures that $\mathcal{U}_{A^R}$ is not empty. The solution to that linear program also provides the necessary and possible assignments of alternative a depending on whether its global utility is no less than that of another alternative for all compatible value functions in $\mathcal{U}_{A^R}$ (necessary) or for at least one of them (possible).

We start with a set of alternatives with decisions, A^R, the *reference set*. We note that $A^R \subseteq A$ where $A = \{a_1, a_2, \ldots, a_n\}$ also includes alternatives without decisions.[31] Each alternative in A was previously evaluated over a set of m criteria: $g_1, g_2, \ldots, g_m$. We let X_j be the resulting vector of assessments on criterion j for the n alternatives and $x_j^1, x_j^2, \ldots, x_j^n$ be those assessments sorted in increasing order.

Drugs included in the reference set have already been classified into q ordered classes, $C_1, C_2, \ldots, C_q$, where C_2 is preferred to C_1 and so on. In our context, we have $q = 3$ classes defined as "full approval" (C_3), "conditional approval" (C_2), and "rejection" (C_1). Rather than using a direct procedure to obtain a global utility function for a given alternative, the preference disaggregation approach uses a regression-based technique to infer the function for that alternative. We assume the function to be additive and to exhibit the following structure:

$$U(a) = \sum_{j=1}^{m} u_j[g_j(a)] \in [0, 1] \tag{1}$$

where $U(a)$ is the global utility of alternative $a \in A$, and $u_j[g_j(a)]$ is the marginal utility of alternative $a \in A$ on evaluation criterion g_j. The global utility $U(a)$, scaled between 0 and 1, represents the evaluation of drug a over all m criteria. A drug will be classified into one of the predefined classes based on the resulting $U(a)$ and a set of utility thresholds $t = [t_1, t_2, \ldots, t_{q-1}]$. Those thresholds, determined optimally, partition the q classes according to the following rule:

$$\left.\begin{array}{rccc} & U(a) & < t_1 & \Rightarrow & a \in C_1 \\ t_k \leq & U(a) & < t_{k+1} & \Rightarrow & a \in C_{k+1},\ k = 1, \ldots, q-2 \\ t_{q-1} \leq & U(a) & & \Rightarrow & a \in C_q \end{array}\right\} \tag{2}$$

The solution to the following linear program yields the marginal utilities and thresholds.

$$Max \quad \epsilon \tag{3}$$

subject to

$$U(a) \geq U(b) + \epsilon \quad \forall a, b \in A^R : C(a) \geq C(b) \tag{4}$$

$$u_j\left(x_j^k\right) - u_j\left(x_j^{k-1}\right) \geq 0 \quad j = 1, 2, \ldots, m,\ k = 2, \ldots, n \tag{5}$$

$$u_j\left(x_j^1\right) = 0 \quad j = 1, 2, \ldots, m \tag{6}$$

$$\sum_{j=1}^{m} u_j\left(x_j^n\right) = 1 \tag{7}$$

$$U(a) \leq t_1 \quad \forall a \in A^R : C(a) = 1 \tag{8}$$

$$t_k \leq U(a) \leq t_{k+1} \quad \forall a \in A^R : C(a) = k, \quad k = 1, \ldots, q-2 \tag{9}$$

$$t_{q-1} \leq U(a) \leq t_{q-1} \quad \forall a \in A^R : C(a) = q \tag{10}$$

$$t_1 \leq \epsilon \tag{11}$$

$$t_k - t_{k-1} \geq \epsilon \quad k = 2, \ldots, q-1 \tag{12}$$

$$t_{q-1} \geq 1 - \epsilon \tag{13}$$

We note that for the objective function (3), a positive optimal value $\epsilon^* > 0$ ensures that $\mathcal{U}_{A^R}$ is not empty. Constraints (4) ensure that the resulting global value function is compatible with all pairs of decisions in set A^R. For instance, alternative a is assigned to a class that is no worse than the class to which alternative b is assigned if and only if the global utility of alternative a is no less than that of alternative b. Constraints (5) ensure the monotonicity of the resulting marginal utilities. Constraints (6) and (7) are normalization constraints that bound marginal utility values between 0 and 1. Constraints (8) to (10) ensure that thresholds cluster alternatives in increasing order of classes. Given a positive optimal value $\epsilon^* > 0$, constraints (11) to (13) make the optimal thresholds distinct and within the $[0, 1]$ interval.

There are other solution approaches to determining value functions. Greco et al.[32] introduced the concept of a representative value function from a set of compatible value functions to make robust sorting recommendations. They obtain this function by maximizing the differences between utilities of alternatives across classes while minimizing differences of utilities within classes.

We implemented the UTADIS^{GMS} method in R statistical software using the Rorutadis library,[33] which is an implementation of the method with a set of value functions compatible with the decision making bodies' preference information. Skedgel et al.[34] developed a data set of pCODR submissions reviewed between 2011 and 2017. With their permission, we use these data as input in the UTADIS^{GMS} model to infer the global utility function and determine the classification thresholds. Subsequently, we conduct scenario analysis by sampling from a set of compatible value functions to predict the assignment of submissions without recommendations (i.e., potential new drugs). Table 1 describes this process.

The data initially included 94 drug submissions with recommendations, 10 criteria, and the full, conditional, or rejection decision for each submission. Three records were removed from the 94 because of missing data. In the data consistency step, UTADIS removed 26 alternatives from the reference set to infer the utility model. An inconsistency occurs when, for a pair of alternatives, the decision of one conflicts with the other, given the criteria assessment values for each. As indicated in Table 1, the program identifies a minimal subset of alternatives that need to be removed so that the set of compatible pairs is free of conflicts and includes a maximum number of consistent alternatives.[35] We included the details of those 26 pulled alternatives in the supplemental material. The resulting reference set had 65 submissions with 11 full approval recommendations, 39 conditional approval recommendations, and 15 rejection recommendations. A drug may appear more than once if it was initially rejected and then resubmitted or if it was resubmitted for a different tumor type or indication.

The criteria are summarized in Table 2. We note that the relative survival gain is a ratio of the survival outcome measure (e.g., progression free, overall survival) of treatment over that of a comparator. If there were no comparators, the ratio was set to 1. The quality of clinical evidence criterion is a binary variable indicating that the submission was based on a 2-arm double-blind randomized phase III clinical trial. The severity of adverse reactions (compared to existing treatments) is also binary, with 1 indicating lower adverse reactions. ICER is a continuous scale variable. The quality of the ICER can be high (= 1) or low (= 0) and indicates confidence in the ICER estimate. The unmet needs criterion is equal to 1 if there are no alternatives. The budget impact is another categorical variable that can assume a high (= 1) or low (= 0) value. To make our model parsimonious and our results statistically robust, we excluded the overall survival flag, infrastructure, and type of drug attributes, since those were correlated with the 7 retained criteria or were not sufficiently discriminating in terms of preferences.

Results

Figure 1 reports the output of the UTADIS^{GMS} model with the global utility values for each drug submission, sorted in decreasing order. The 2 dashed vertical lines indicate the optimal threshold values between the

Table 1 Description of the Method

Phase	Step	Output	Notes
Structuring the problem	Identification of criteria (Skedgel et al.[34])	Of the 10 criteria identified, we retained 7 in our analysis.	We kept criteria that are most relevant (i.e., independent and discriminating in terms of preferences).
	Tabulation of alternatives (Skedgel et al.[34])	There were 91 alternatives with decisions in their final sample size.	
	Measurement of performance (Skedgel et al.[34])	The performance matrix includes criteria assessments and decisions for each of the 91 alternatives.	See Input & Output tab in the supplemental material spreadsheet.
	Scenario development	We developed 3 scenarios (alternatives without a decision) and appended their criteria assessments to the set of alternatives.	See Input & Output tab in the supplemental material spreadsheet.
Applying UTADIS^{GMS}	Data consistency	Inconsistencies are identified.	Rorutadis identifies pairs of alternatives with conflicts. See examples in the Inconsistencies tab in the supplemental material.
		Alternatives causing inconsistencies are removed.	Rorutadis ensures that the set of removed alternatives is of minimal size.
	Generation of results	The output includes: 1. The marginal utility functions for each criterion 2. The thresholds that partition the 3 classes 3. The global utility value and class assignment for all alternatives, including the scenarios	See Input & Output tab in the supplemental material spreadsheet.
Analysis and model validation	Cross-validation	Thresholds and classification for each testing sample	See Cross-validation tab in the supplemental material spreadsheet.
	Robustness of scenario results	Percentage of value functions for which scenarios remain in the same class assignment	Results from Rorutadis's *CalculateStochasticResults* function are reported in the Results section.

Table 2 Evaluation Criteria and Scenario Values

Evaluation Criteria	Definition	Scenario 1	Scenario 2	Scenario 3
Relative survival gain	Survival gain v. comparator (continuous, range = [0.8, 5.3])	1.1	1.5	1.0
Quality of clinical evidence	Quality of clinical study with respect to phase of the random clinical trial and results (binary, 1 = high quality, 0 = low quality)	1	1	0
Severity of adverse reactions	Severity of adverse reactions compared to existing treatment (binary, 1 = lower adverse reactions, 0 = higher adverse reactions)	1	0	1
ICER	Incremental cost-effectiveness ratio ($K, continuous, range = [18, 461])	75	150	200
ICER quality	Uncertainty of ICER, sensitivity analysis (binary, 1 = high uncertainty, 0 = low uncertainty)	0	0	1
Unmet needs	(binary: 1 = existence of alternatives, 0 = absence of alternatives)	1	1	0
Budget impact	Estimated impact based on the patient population size and available treatments (binary, 1 = high impact, 0 = low impact)	1	0	0
Overall survival flag	Not included			
Infrastructure	Not included			
Type of drug	Not included			

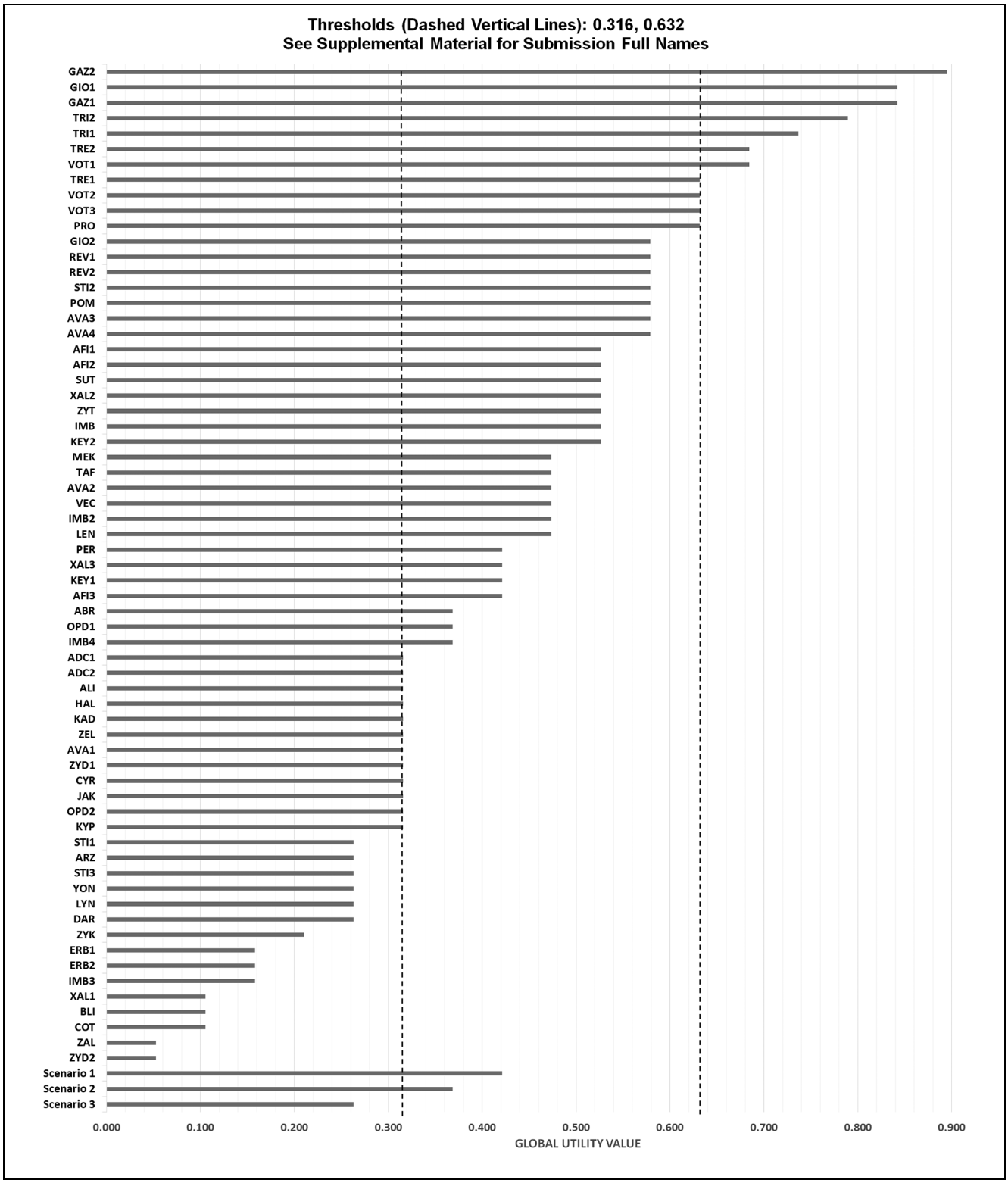

Figure 1 Global utility values and classification of alternatives.

rejected and conditional approval classes (0.316) and between the conditional and full approval classes (0.632).

The last 3 rows of Figure 1 report the global utility values and the classes of 3 scenarios (potential new drugs) that we model with the vectors of criteria assessments in Table 2. We label the hypothetical drugs "incremental" (scenario 1), "drastic" (scenario 2), and "orphan" (scenario 3). Scenario 1 models a potential incremental drug, with a relative survival gain of 1.1 (or 10% over an existing comparator). We further assume that it has a high quality of clinical evidence, that it causes lower adverse reactions, and that alternative options already exist. We set its ICER at $75,000, which corresponds to the lower quartile values for the 65 submissions, and a high-budget impact that would result from high demand. Rorutadis estimates the global utility value of this potential alternative at 0.421. Since it lies between the 2 thresholds (0.316, 0.632), it falls into the conditional approval class.

Scenario 2 models a potential drastic drug that is more efficacious with a relative survival gain of 1.5. We assume it exhibits a higher ICER than the incremental drug ($150,000) and higher adverse reactions compared to its existing alternatives. Contrary to the incremental drug in scenario 1, we set the budget impact at "low" based on lower expected demand. Its global utility returned at 0.368 and is thus also categorized in the conditional approval class.

Scenario 3 models a potential orphan drug. We set its relative survival gain at 1.0, since there is no comparator, and the quality of clinical evidence at "low." We assume it is at phase II of a single-arm clinical trial and that it will cover unmet therapeutic needs. We further assume low adverse reactions, no alternative treatment, and a low budget impact due to low demand. We set its ICER at $200,000, which corresponds to the higher quartile values of the reference set. We also assume its ICER exhibits high uncertainty, which is typical of orphan drugs. The resulting global utility value is 0.263, which puts it in the rejected class.

In a disaggregation approach, fit statistics and related error measures are moderately useful for validation purposes. What is central to validation is interaction with the decision makers.[17,18] Since it is their decision process that is being represented, validation consists of informing them of the model output. They can analyze this information to review or adjust the decision criteria and reference set.

We nevertheless performed cross-validation by applying the model 5 times with distinct testing samples of $n = 5$ alternatives pulled from the reference set and another 5 times with testing samples of $n = 10$ alternatives. Details are reported in the supplemental material. In the $n = 5$ tests, we noted that only 1 alternative in one sample was categorized differently. In the $n = 10$ tests, we noted 2 samples with 4 alternatives categorized differently, 2 samples with 2 alternatives categorized differently, and 1 with a single alternative categorized differently. Such results illustrate the impact of eroding information from the increasing test samples on the model output. Less data implies a higher propensity of bias and a diminished learning capacity of the model. This further justifies interaction with the committee to determine the appropriate size of the reference set for this context.

We performed a validation exercise more in line with a disaggregation approach by applying the *CalculateStochasticResults* Rorutadis function. It tested the 3 scenarios against a randomly generated sample of 100 representative functions. It assigned the incremental drug to the conditional approval class for 89% of the representative functions and to the full approval class for 11%. It assigned the drastic drug to the conditional approval class for 100% of the functions and the orphan drug to the rejected class for 100% of the functions. These percentages show that the model is robust since the assignments are almost the same for the sample of 100 representative functions.

Discussion

In this article, we proposed a methodological framework of MCDA to investigate the potentials and applicability of a preference disaggregation approach to drug formulary design. The main advantage of this approach is that it infers preferences from the data as opposed to other MCDA methods that rely on direct assessments of model parameters and consequently on a greater cognitive effort from the decision making bodies.[31] In addition to objectively sorting and determining thresholds for the inclusion of existing drugs, this method can be applied to a new drug using estimates or targets for the criteria assessments to help predict its recommendation outcome. This is useful information that can assist review boards in their decision process. It can also help research and development divisions of companies to prioritize resources and innovation capabilities.

The method having been successful in numerous fields, most limitations of this work deal with the formulary application. We included criteria and alternatives based on the Skedgel et al.[34] analysis. We could add criteria that explicitly assess clinical judgment, demand, patient-based values, or adoption feasibility. We could also expand the binary criteria scales to reflect measurements

that are more accurate. For instance, the range for the quality of the clinical evidence criterion could be widened in light of more precise information. In addition, although the committee reviews dozens of submissions each year, the model does not consider the timing of the recommendations. Thus, as new submissions and decisions become available, we could update the reference set to reflect a more current or precise set of alternatives that captures the evolving preferences of the committee. For instance, we could truncate older submissions from the set or partition it into subsets defined by tumor type or indication.

In the MCDA literature, disaggregation methods are all based on additive utility functions. When the criteria are independent, this is sufficient to capture the decision makers' preferences.[22] Independence is a reasonable assumption in our pCODR context, as shown in Table 2. In more complex decision making settings, an additive utility function may not properly capture interactions between criteria. A more general form could possibly be developed and tested to model preferences.

Our model proved robust based on the results obtained from applying the *CalculateStochasticResults* Rorutadis function. In MCDA, robustness refers to the analysis of imperfections of the model. Assessing robustness consists of verifying whether the provided conclusions match those from a randomly generated set of compatible utility functions. Experimental work shows that the robustness of an additive model depends on both the number of criteria and the size of the reference set.[36]

The approach allows for the identification of inconsistent decisions. This resulted in the removal of 26 alternatives from the reference set to infer the utility model. Inconsistencies can be explained by many factors. In decision making, criteria not explicitly stated can nonetheless influence outcomes. Also, perceptions and preferences of the decision making body may change over time and context. For example, assessment of budget impact may depend on the economic context in which the review occurs. Moreover, committee membership turnover may influence the way it processes information and makes recommendations. This limits the overall "memory" of the committee and does not ensure that the recommendation of a new submission is consistent with previous ones. In addition, committee members, based on their own field of expertise, may favor submissions that affect some patient groups, conditions, or tumor types over others.[34] To control inconsistencies, the disaggregation approach calls for continuous interaction and feedback with the committee.[18] This exercise helps calibrate the model and ultimately lowers the number of inconsistent decisions.

Disaggregation analysis does not replace but rather facilitates the decision support process. It helps decision makers gain insight into the problem's alternatives and criteria, the modeling process, and the interpretation of its output. The inferred model provides a starting basis for the decision making support process. If the model output agrees with the preferential system of the decision makers, then the model can be directly applied to new decision instances. Otherwise, the decision makers can revise their criteria and reference set or provide recommendations about the calibration of the model. This approach instills an efficient, quick, and reliable way to support the review process that ensures inclusion of relevant criteria and consistency among decisions.

In a wider analysis, more therapeutic conditions can be modeled in the set of alternatives of the formulary. Moreover, formularies can be multitiered, with each tier defined by a cost-sharing policy, which is typical of private insurance plans. Insurers and pharmacy benefit managers also design formularies for a multiplicity of plans. Thus, formulary candidates can be modeled as alternatives by developing appropriate sets of criteria. The methodology above could then be extended and applied to glean strategic value of candidate formularies, not only to insurers but also to employers or other clients who, in turn, provide coverage options to their workforce or constituents.

This article contributes to an ongoing discussion on MCDA use in HTA decisions,[31,37,38] specifically the critical value of proven and tested methodology in evaluating treatment options. We focused on recent extensions to the UTADIS method and expanded it to model the reimbursement recommendation decision of new or existing prescription drugs. We applied it using data on Canadian oncology drugs reviewed between 2011 and 2017. By associating the input and requirements of the method with the information germane to the offerings, we obtained valuable information on their potential outcome in terms of global utility value and class. We also developed scenarios that illustrate how the method can be applied for predictive purposes.

Acknowledgments

The authors thank Krzysztof Ciomek, PhD, from Poznan University of Technology for his valuable advice and Christopher Skedgel, PhD, from the University of East Anglia for the permission to use the data set of decision summaries.

ORCID iD

Gilles Reinhardt https://orcid.org/0000-0001-9810-9503

Supplemental Material

Supplementary material for this article is available on the *Medical Decision Making* Web site at http://journals.sagepub.com/home/mdm.

References

1. Canadian Agency for Drugs and Technologies in Health (CADTH.ca). Process in brief. Available from: https://www.cadth.ca/about-cadth/what-we-do/products-services/cdr/common-drug-review-submissions/process-in-brief
2. CADTH pan-Canadian Oncology Review. Find a review (pCODR). Available from:https://www.cadth.ca/pcodr/find-a-review
3. Stinnett AA, Paltiel AD. Mathematical programming for the efficient allocation of health care resources. *J Health Econ*. 1996;15(5):641–53.
4. Weintraub WS, Cohen DJ. The limits of cost-effectiveness analysis. *Circ Cardiovasc Qual Outcomes*. 2009;2(1):55–8.
5. Olmstead T, Zeckhauser R. The menu-setting problem and subsidized prices: drug formulary illustration. *J Health Econ*. 1999;18(5):523–50.
6. Truong V-A. Optimal selection of medical formularies. *J Revenue Pricing Manag*. 2014;13(2):113–32.
7. Roy B. *Multicriteria Methodology for Decision Aiding*. Boston, MA: Springer Science & Business Media; 2013.
8. Marsh K, Lanitis T, Neasham D, Orfanos P, Caro J. Assessing the value of healthcare interventions using multi-criteria decision analysis: a review of the literature. *Pharmacoeconomics*. 2014;32(4):345–65.
9. Thokala P, Devlin N, Marsh K, et al. Multiple criteria decision analysis for health care decision making—an introduction: report 1 of the ISPOR MCDA Emerging Good Practices Task Force. *Value Health*. 2016;19(1):1–13.
10. Marsh K, IJzerman M, Thokala P, et al. Multiple criteria decision analysis for health care decision making—emerging good practices: report 2 of the ISPOR MCDA Emerging Good Practices Task Force. *Value Health*. 2016;19(2):125–37.
11. Youngkong S, Baltussen R, Tantivess S, Mohara A, Teerawattananon Y. Multicriteria decision analysis for including health interventions in the universal health coverage benefit package in Thailand. *Value Health*. 2012;15(6):961–70.
12. Endrei D, Molics B, Ágoston I. Multicriteria decision analysis in the reimbursement of new medical technologies: real-world experiences from Hungary. *Value Health*. 2014;17(4):487–9.
13. Thokala P, Duenas A. Multiple criteria decision analysis for health technology assessment. *Value Health*. 2012;15(8):1172–81.
14. Siskos Y, Spyridakos A. Intelligent multicriteria decision support: overview and perspectives. *Eur J Oper Res*. 1999;113(2):236–46.
15. Marsh K, Sculpher M, Caro JJ, Tervonen T. The use of MCDA in HTA: great potential, but more effort needed. *Value Health*. 2018;21(4):394–7.
16. Walker A. Challenges in using MCDA for reimbursement decisions on new medicines? *Value Health J Int Soc Pharmacoeconomics Outcomes Res*. 2016;19(2):123–4.
17. Doumpos M, Zopounidis C. Preference disaggregation and statistical learning for multicriteria decision support: a review. *Eur J Oper Res*. 2011;209:203–14.
18. Corrente S, Greco S, Kadziński M, Słowiński R. Robust ordinal regression in preference learning and ranking. *Mach Learn*. 2013;93(2):381–422.
19. 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. *Manag Sci*. 2008;54(7):1336–49.
20. Greco S, Mousseau V, Słowiński R. Multiple criteria sorting with a set of additive value functions. *Eur J Oper Res*. 2010;207(3):1455–70.
21. Devaud JM, Groussaud G, Jacquet-Lagreze E. UTADIS: Une méthode de construction de fonctions d'utilités additives rendant compte de jugements globaux. Paper presented at: European Working Group on Multicriteria Decision Aid; October 1980; Bochum, Germany.
22. Jacquet-Lagreze E, Siskos J. Assessing a set of additive utility functions for multicriteria decision-making, the UTA method. *Eur J Oper Res*. 1982;10(2):151–64.
23. Kosmidou K, Zopounidis C. Predicting US commercial bank failures via a multicriteria approach. *Int J Risk Assess Manag*. 2008;9(1–2):26–43.
24. Gaganis C, Pasiouras F, Zopounidis C. A multicriteria decision framework for measuring banks' soundness around the world. *J Multicriteria Decis Anal*. 2006;14:103–11.
25. Pendaraki K, Zopounidis C, Doumpos M. On the construction of mutual fund portfolios: a multicriteria methodology and an application to the Greek market of equity mutual funds. *Eur J Oper Res*. 2005;163(2):462–81.
26. Voulgaris F, Doumpos M, Zopounidis C. On the evaluation of Greek industrial SME's performance via multicriteria analysis of financial ratios. *Small Bus Econ*. 2000;15(2):127–36.
27. Grigoroudis E, Orfanoudaki E, Zopounidis C. Strategic performance measurement in a healthcare organisation: a multiple criteria approach based on balanced scorecard. *Omega*. 2012;40(1):104–19.
28. Siskos Y, Yannacopoulos Y. UTASTAR: an ordinal regression method for building additive value functions. *Investig Oper*. 1985;5(1):39–53.
29. Zopounidis C, Doumpos M. PREFDIS: a multicriteria decision support system for sorting decision problems. *Comput Oper Res*. 2000;27(7):779–97.
30. Esmaelian M, Shahmoradi H, Nemati F. A multi criteria classification method. *Curr Future Dev Artif Intell*. 2017;1:214–67.
31. Jacquet-Lagrèze E, Siskos Y. Preference disaggregation: 20 years of MCDA experience. *Eur J Oper Res*. 2001;130(2):233–45.

32. Greco S, Kadziński M, SŁowiński R. Selection of a representative value function in robust multiple criteria sorting. *Comput Oper Res*. 2011;38(11):1620–37.
33. Package "rorutadis." Robust Ordinal Regression UTADIS. January 17, 2017. Available from: https://cran.r-project.org/web/packages/rorutadis/rorutadis.pdf
34. Skedgel C, Wranik D, Hu M. The relative importance of clinical, economic, patient values and feasibility criteria in cancer drug reimbursement in Canada: a revealed preferences analysis of recommendations of the Pan-Canadian Oncology Drug Review 2011–2017. *Pharmacoeconomics*. 2018;36(4):467–75.
35. Kadziński M, Cinelli M, Ciomek K, et al. Co-constructive development of a green chemistry-based model for the assessment of nanoparticles synthesis. *Eur J Oper Res*. 2018;264(2):472–90.
36. Kadziński M, Ghaderi M, Wąsikowski J, Agell N. Expressiveness and robustness measures for the evaluation of an additive value function in multiple criteria preference disaggregation methods: an experimental analysis. *Comput Oper Res*. 2017;87:146–64.
37. Angelis A, Kanavos P. Multiple criteria decision analysis (MCDA) for evaluating new medicines in health technology assessment and beyond: the advance value framework. *Soc Sci Med*. 2017;188:137–56.
38. Hall W, Mitton C, Levy A. Realizing the "great potential" of MCDA in HTA. *Value Health*. 2018;21:1355–6.