Original Article

Value Assessment of Oncology Drugs Using a Weighted Criterion-Based Approach

Doreen A. Ezeife, MD [1]; Francois Dionne, PhD[2]; Aline Fusco Fares, MD[3]; Ellen Laura Rose Cusano, MD[1]; Rouhi Fazelzad, BSc, MISt[4]; Wenzie Ng, BSc, MPharm, RPh[3]; Don Husereau, BSc Pharm, MSc[5]; Farzad Ali, BPharm, MSc[6]; Christina Sit, MSc[7]; Barry Stein, B.Com BCL, LLB[8]; Jennifer H. Law, MSc[3]; Lisa Le, MSc[3]; Peter Michael Ellis, MD, MMed, PhD[9]; Scott Berry, MD[10]; Stuart Peacock, PhD[11]; Craig Mitton, PhD[12]; Craig C. Earle, MD[13]; Kelvin K. W. Chan, MD, PhD [13]; and Natasha B. Leighl, MD, MMSc[3]

BACKGROUND: Globally, the rising cost of anticancer therapy has motivated efforts to quantify the overall value of new cancer treatments. Multicriteria decision analysis offers a novel approach to incorporate multiple criteria and perspectives into value assessment. **METHODS:** The authors recruited a diverse, multistakeholder group who identified and weighted key criteria to establish the drug assessment framework (DAF). Construct validity assessed the degree to which DAF scores were associated with past pan-Canadian Oncology Drug Review (pCODR) funding recommendations and European Society for Medical Oncology Magnitude of Clinical Benefit Scale (ESMO-MCBS; version 1.1) scores. **RESULTS:** The final DAF included 10 criteria: overall survival, progression-free survival, response rate, quality of life, toxicity, unmet need, equity, feasibility, disease severity, and caregiver well-being. The first 5 clinical benefit criteria represent approximately 64% of the total weight. DAF scores ranged from 0 to 300, reflecting both the expected impact of the drug and the quality of supporting evidence. When the DAF was applied to the last 60 drugs (with reviewers blinded) reviewed by pCODR (2015-2018), those drugs with positive pCODR funding recommendations were found to have higher DAF scores compared with drugs not recommended (103 vs 63; Student *t* test $P = .0007$). DAF clinical benefit criteria mildly correlated with ESMO-MCBS scores (correlation coefficient, 0.33; 95% CI, 0.009-0.59). Sensitivity analyses that varied the criteria scores did not change the results. **CONCLUSIONS:** Using a structured and explicit approach, a criterion-based valuation framework was designed to provide a transparent and consistent method with which to value and prioritize cancer drugs to facilitate the delivery of affordable cancer care. ***Cancer*** **2020;126:1530-1540**.

KEYWORDS: anticancer therapy, multicriteria decision analysis, value assessment, value assessment framework.

INTRODUCTION

In recent years, the number of available oncology treatments has significantly increased and these treatments have become more costly, prompting a growing need for sustainable drug funding decision-making processes. In the United States, Canada, and many other countries around the world, cancer drug budgets are increasing at a much faster rate than the gross domestic product per capita.[1] From a societal perspective, there are growing concerns that many of the expensive new cancer medicines that are being developed offer limited meaningful benefit to patients.[2,3] Cost-constrained health care systems are struggling to provide timely access to new cancer treatments, and high drug acquisition costs inevitably result in significant delays in uptake.[4,5]

Economically diverse countries face inconsistencies in access that have motivated major oncology societies to develop standardized tools for value assessment.[6-9] The American Society of Clinical Oncology's (ASCO) framework evaluates new cancer drugs at the individual patient level to inform overall value for clinical decision making. The European Society for Medical Oncology Magnitude of Clinical Benefit Scale (ESMO-MCBS) assesses clinical benefit and was proposed

Corresponding Author: Doreen A. Ezeife, MD, Tom Baker Cancer Centre, University of Calgary, 1331 29 St NW, Calgary, AB T2N 4N2 Canada (doreen.ezeife@albertahealthservices.ca).

[1]Department of Oncology, Tom Baker Cancer Centre, Calgary, Alberta, Canada; [2]Prioritize Consulting Ltd, Vancouver, British Columbia, Canada; [3]Department of Medical Oncology and Hematology, Princess Margaret Cancer Centre, Toronto, Ontario, Canada; [4]Department of Medical Oncology and Hematology, Library and Information Services, Princess Margaret Cancer Centre, Toronto, Ontario, Canada; [5]School of Epidemiology and Public Health, University of Ottawa, Ottawa, Ontario, Canada; [6]Pfizer Inc, New York, New York; [7]Lung Cancer Canada, Toronto, Ontario, Canada; [8]Colorectal Cancer Canada, Montreal, Quebec, Canada; [9]Juravinski Cancer Centre, Hamilton, Ontario, Canada; [10]Cancer Centre of Southeastern Ontario, Kingston, Ontario, Canada; [11]British Columbia Cancer Agency, Vancouver, British Columbia, Canada; [12]The University of British Columbia, Vancouver, British Columbia, Canada; [13]Sunnybrook Odette Cancer Centre, Toronto, Ontario, Canada

The last 2 authors are co-senior authors.

We thank Mr. Khizar Karim, Mr. Malcolm Ryan, and the patients who contributed to this work. We also acknowledge the Canadian Association of Medical Oncologists, the Canadian Institutes of Health Research, and the Department of Medical Oncology at Princess Margaret Cancer Centre for providing funding for this work.

Additional supporting information may be found in the online version of this article.

DOI: 10.1002/cncr.32639, **Received:** May 31, 2019; **Revised:** October 9, 2019; **Accepted:** October 24, 2019, **Published online** December 20, 2019 in Wiley Online Library (wileyonlinelibrary.com)

by ESMO to guide drug approval decisions, although the European Medicines Agency and individual European countries make their own independent funding decisions. In Canada, the pan-Canadian Oncology Drug Review (pCODR) provides national evidence-based funding recommendations using a deliberative framework that considers additional factors beyond the traditional pillars of clinical benefit and cost.[10] Although the dimensions of value considered by pCODR are clearly defined, to the best of our knowledge there is no explicit weighting scheme for these criteria.

Historically, explicit decision-making processes have been promoted as a way to reinforce public trust in the fair allocation of limited resources.[11,12] When society is considering the adoption of new cancer medicines, value assessment often requires a more comprehensive and systematic consideration of multiple perspectives than any of the current value assessment tools offer. For many new cancer medicines, the cost per life-year gained now routinely exceeds US $150,000,[13,14] and thus cancer drug funding decisions are increasingly contentious. Understandably, societal demands for legitimacy, consistency, and accountability for these decisions are growing.

For complex health care decisions, decision analysis can provide a structured, explicit approach to facilitate decisions involving trade-offs between possible benefits and harms.[15,16] A type of decision analysis called multicriteria decision analysis (MCDA) has been gaining popularity in health care and oncology decision making. In the United Kingdom, MCDA-based decision-making frameworks have been developed to facilitate orphan drug reimbursement decisions.[17] The Institute of Medicine in the United States also has recommended a prioritization framework for vaccines, using MCDA principles.[18] The International Society for Pharmacoeconomics and Outcomes Research complemented the recent surge of interest in MCDA by establishing best practice guidelines to facilitate the incorporation of MCDA into health care decision making.[15,16]

The MCDA approach summarizes varying and often conflicting stakeholder opinions to reach transparent and consistent decisions. Using MCDA, a broader set of values are captured by identifying the relevant criteria and assigning weights to these criteria. We applied MCDA methods to the development of a value assessment framework for oncology drugs. In addition, we validated the framework by correlating framework scores with pCODR funding recommendations and ESMO-MCBS thresholds for meaningful benefit.

MATERIALS AND METHODS

Development of the Framework

Stakeholder selection

Stakeholders from diverse perspectives of society were recruited to participate in the development of the drug assessment framework (DAF). This study was approved by the University of Toronto Research Ethics Board and the Sunnybrook Odette Cancer Centre Research Ethics Board. The framework was developed with close adherence to the International Society for Pharmacoeconomics and Outcomes Research MCDA Good Practice guidelines.[15,16] Stakeholders included patients (2 individuals), public members (2 individuals), patient advocacy group leaders (2 individuals), pharmacists (2 individuals), industry representatives (1 individual), oncologists (6 individuals), ethicists (1 individual), health economists (3 individuals), pCODR members (3 individuals), cancer agency members (2 individuals), and a Ministry of Health government representative. Some stakeholders shared multiple perspectives.

Identifying criteria

Criteria initially were identified through a systematic search of the MEDLINE, Excerpta Medica database (EMBASE), and Epub Ahead Ovid of Print databases as shown in Table 1[6,7,9,10,19-25] (Fig. 1) (see Supporting Information Materials). These criteria identified the unique domains of drug value and were framed into an initial set of criteria believed to be relevant to cancer drug funding decisions. In telephone consultation with stakeholders, caregiver well-being was added to the initial list. Stakeholders then were asked to complete an online survey (see Supporting Information Materials) that asked the following question: for cancer drug funding decisions, how important is consideration of _____[insert criteria]_____. In response to this question, stakeholders scored each criterion on a scale from 0 to 5, with 0 being the least important and 5 being the most important for cancer drug funding decisions. On the basis of survey responses, the innovation criterion was eliminated. The final set of 10 criteria were reviewed for completeness, nonredundancy, nonoverlap, and preference independence.

Weighting criteria and calculating the total score

The ranking method was used to assign weights to the criteria. The ranking method has been used in previous health care studies,[26] and provides the level of precision required in MCDA weighting methods while minimizing stakeholder cognitive burden and maintaining accuracy.

TABLE 1. Criteria Identification From Prior Studies

Study	Study Objectives	Criteria Used
Angelis 2017[19]	To assess the value of second-line biological treatments in metastatic colorectal cancer after prior oxaliplatin-based chemotherapy using MCDA	OS, adverse events, HRQOL, medical costs impact, posology, PFS, marketing authorization, innovation
American Society of Clinical Oncology 1996[20]	To define the outcomes of cancer treatment that should be considered for technology assessment and cancer treatment guidelines	Survival (ie, OS, PFS, EFS, and DFS), QOL, toxicity, measures of cancer response (ie, RR, biomarkers, cancer-induced abnormalities in common blood tests), cost-effectiveness
Browman 2008[21]	To prioritize new oncology drugs seeking funding in the Canadian provincial (Alberta) cancer system	Clinical benefit, strength of evidence, consistency of results (ie, magnitude and direction of effect), clinical impact (ie, size and direction of effect across studies), appropriateness of measures used to assess outcomes, toxicity/convenience, availability of alternatives
Cherny 2015[9]	To assess the magnitude of clinical benefit for cancer medicines	Clinical benefit (OS, PFS, DFS, and EFS), QOL, toxicity
Cheung 2016[10]	To compare the perspectives of value that exist for the various HTA frameworks (Europe vs Canada)	Clinical effectiveness, safety, burden of illness, unmet need, alignment with patient values, cost-effectiveness, economic feasibility (budget impact), organizational feasibility and Cherny 2015 criteria[9] (see above)
Kwon 2017[22]	To elicit societal preferences of reimbursement decision criteria for anticancer drugs using MCDA	Disease severity, disease population size, therapeutic target for pediatrics, unmet needs, innovation, clinical benefit, cost-effectiveness, budget impact
Leung 2017[23]	To identify oncology treatments suitable for retrospective outcomes analysis to inform continued funding decisions in British Columbia, Canada	Strength and consistency of evidence, clinical efficacy, safety, ICER, feasibility, interprovincial and national funding decision(s)
MacLeod 2016[24]	To identify and compare funding preferences of patients with cancer and the general public against the criteria used by payers making cancer drug funding decisions	Clinical efficacy, safety, QOL, societal benefit, disease severity, unmet need, cost-effectiveness, budget impact, feasibility, strength of evidence, equity, innovation
Regier 2014[25]	To examine priority setting for new cancer drugs	Clinical benefit, magnitude of benefit, quality of evidence, availability of alternatives, cost, duration of treatment, budget impact
Schnipper 2015[6] and Schnipper 2016[7]	To develop a standardized approach to assist physicians and patients in assessing the value of new anticancer drugs	Clinical benefit (OS, PFS, and RR), toxicity, duration of benefits, symptom palliation, QOL, treatment-free interval

Abbreviations: DFS, disease-free survival; EFS, event-free survival; HRQOL, health-related quality of life; HTA, health technology assessment; ICER, incremental cost-effectiveness ratio; MCDA, multicriteria decision analysis; OS, overall survival; PFS, progression-free survival; QOL, quality of life; RR, response rate.

Stakeholders were asked to list the 3 most important criteria and the 3 least important criteria. Criteria ranking was aggregated according to the following MCDA ranking principles:

- Criteria weights total 100.
- Criteria were grouped into 3 categories: 1) highly ranked criteria (assigned a starting weight of 15); 2) medium ranked criteria (assigned a starting weight of 10); and 3) lowest ranked criteria (assigned a starting weight of 5).
- The final weight for each criterion was calculated by adjusting the starting weight by 0 to 4 points based on the absolute frequency of high or low ranking.

Aggregate stakeholder rankings of the criteria can be found in the Supporting Information Materials.

For each criterion, a rating scale was implemented for scoring. The rating scale is a 4-point Likert scale ranging from 0 to 3, representing worst to best performance, respectively. For the clinical benefit criteria, the rating scale was formulated based on previously published work regarding the integration of absolute differences and effect estimates into data interpretation.[9,27] A strength of evidence (SOE) modifier was implemented that modified a criterion's score based on supporting evidence. If strong, moderate, or weak evidence was used, then 0, 1, or 2 points, respectively, were deducted from the criterion score.

Total score contributions were calculated by the sum of each criterion's weight multiplied by its score, as represented by the following additive model:

$$\text{Total Score} = \left(\text{weight}_1 \times \text{score}_1\right) + \left(\text{weight}_2 \times \text{score}_2\right) + \ldots . + \left(\text{weight}_n \times \text{score}_n\right)$$

Validation of the Framework

Face and content validation

The initial framework was reviewed through one-on-one meetings with each stakeholder to establish face and content validity. Stakeholders reviewed the developed framework for consistency, completeness, and relevance to the decision problem. In this iterative face and content validation process, necessary modifications were made to arrive at the final DAF.

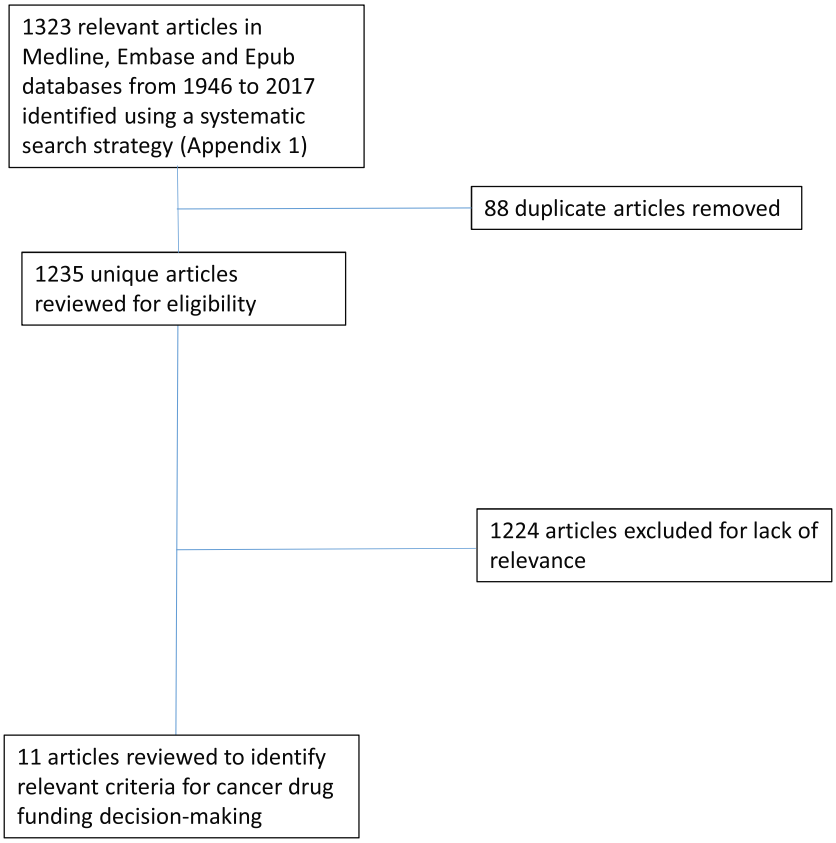

Figure 1. Eligible studies for the multicriteria decision analysis drug assessment framework according to MEDLINE, Excerpta Medica database (EMBASE), Ovid Epub Ahead of Print and In-Process databases, and other nonindexed citations (see Supporting Information Materials).

Construct validation

To ensure construct validity of the framework (ie, that it measures the value of novel cancer drugs), the framework was applied retrospectively to the last 60 drugs reviewed by pCODR between September 2015 and May 2018. Eligible drugs also were scored using the ESMO-MCBS (version 1.1).[9]

For the retrospective assessment, pCODR deliberations and funding recommendations were independently coded by at least 2 blinded scorers (D.A.E., A.F.F., and E.L.R.C.).[28] If there was disagreement, a consensus was reached to determine a final score. During the scoring process, scorers were blinded to the pCODR funding recommendation and ESMO-MCBS scores. The pCODR funding recommendation was recorded independently by a data collector. The ESMO-MCBS scores were obtained from the published scores for those available,[8] and were derived independently by 2 data collectors for those without published scores.

Uncertainty analysis

For the qualitative criteria (quality of life [QOL], unmet need, toxicity, equity, disease severity, and caregiver well-being), a sensitivity analysis was conducted to explore the impact of variability in assigning scores. For each qualitative criterion, one point was deducted from the initially assigned score to represent the lower limit score, and one point was added to represent the upper limit score. The analysis was rerun with the lower limit and upper limit aggregate scores. Scenario analysis was conducted using a subset of the criteria, namely the clinical benefit criteria (QOL, overall survival [OS], progression-free survival [PFS], response rate [RR], and toxicity).

Statistical Analysis

Descriptive statistics summarized DAF scores for the 60 drugs reviewed. Associations between total DAF score, clinical benefit score, or the SOE modifier and pCODR funding recommendation were examined using

			RATING				STRENGTH OF EVIDENCE MODIFIER			
CRITERIA	DEFINITION	WEIGHT	0	1	2	3	No change	-1	-2	
1) Quality of life	Impact of the new treatment on overall sense of well-being as measured by QOL instruments or feedback from patients, compared to existing treatments.	19	No evidence of impact on QOL	Minimal QOL improvement reported (using validated scale)	Moderate QOL improvement reported (using validated scale)	Significant QOL improvement reported (using validated scale)	Strong evidence	Medium strength of evidence	Weak evidence	Level 1=high level of confidence in the evidence (eg. High quality RCT(s) with convincing, consistent results and free from bias) Level 2=moderate level of confidence (eg. RCT data with some identified bias) Level 3=low level of confidence (eg. Significant concerns about the validity of the available evidence) Note: a new treatment cannot be assessed until the corresponding clinical trial has been published.
2) Overall survival (OS), HR=hazard ratio	A new treatment's ability to increase survival time (from treatment start to death from any cause), compared to existing treatments.	15	OS: HR non-inferior or no difference in median OS	OS: HR>0.8 or < 30% improvement in median OS	OS: HR 0.6-0.8 or 31-60% improvement in median OS	OS: HR<0.6 or > 61% improvement in median OS	Strong evidence	Medium strength of evidence	Weak evidence	For OS and PFS/DFS, use the magnitude of improvement in median survival, if reported. If not reported, use hazard ratio.
3) Unmet clinical need	Existence of other treatments for the underlying condition, for the specific population targeted by the drug. The question to be asked is: in the absence of this drug, do clinicians have any available option for treatment or is their only option to focus on limiting the symptoms	15	Alternative treatment (drug or non-drug) exists for all or almost all needs for this patient population	Alternative treatment (drug or non-drug) exists for most needs for this patient population	Alternative treatment (drug or non-drug) exists for some of the needs of this patient population	These patients currently have none or very limited alternative treatment options.	Strong evidence	Medium strength of evidence	Weak evidence	
4) Progression-free survival or Disease-free survival (DFS), HR=hazard ratio	A new treatment's ability to increase survival time without the cancer getting worse, compared to existing treatments. Disease-free survival: a new treatment's ability to increase time living without disease, compared to existing treatments.	12	No evidence of increase in DFS	PFS/DFS improvement alone HR>0.8 or < 30% improvement in median OS or DFS improvement <1.5%	PFS/DFS improvement alone HR 0.6-0.8 or or 31-60% improvement in median OS or DFS improvement 1.5-5%	PFS/DFS improvement alone HR <0.6 or > 61% improvement in median OS or DFS improvement >5%	Strong evidence	Medium strength of evidence	Weak evidence	
5) Toxicity	Side effects of a new treatment, compared to existing treatments.	10	More substantial safety concerns or significant adverse events reported than with comparator	Safety profile similar to comparator	Fewer safety concerns reported or adverse events reported than with comparator	Reliable evidence of substantial reduction in safety concerns or adverse events reported than with comparator	Strong evidence	Medium strength of evidence	Weak evidence	
6) Response rate	A new treatment's ability to cause the cancer to shrink, compared to existing treatments. Response rate = Percentage of complete + partial response.	8	No evidence of increase in response rate	Minimal impact on response rate (<30%)	Moderate impact on response rate (31-60%)	Significant impact on response rate (>61%)	Strong evidence	Medium strength of evidence	Weak evidence	
7) Equity	Impact of the new treatment on the health of vulnerable or marginalized populations, compared to existing treatments.	6	No impact on disparities	Minimal improvement in the health of vulnerable populations	Moderate improvement in the health of vulnerable populations	Significant improvement in the health of vulnerable populations	Strong evidence	Medium strength of evidence	Weak evidence	
8) Feasibility	The difference in the ease of administration, implementation or accessibility of the new treatment, compared to existing treatments.	6	No impact on the ease of administration	Minimal improvement in the ease of administration	Moderate improvement in the ease of administration	Significant improvement in the ease of administration	Strong evidence	Medium strength of evidence	Weak evidence	
9) Disease severity	Mortality or morbidity rate of the cancer.	5	Cancer associated with **very low** mortality or morbidity	Cancer associated with **low** mortality or morbidity	Cancer associated with **moderate** mortality or morbidity	Cancer associated with **high** mortality or morbidity	Strong evidence	Medium strength of evidence	Weak evidence	
10) Caregiver well-being	The extent to which the new treatment improves caregiver well-being, compared to existing treatments.	4	No evidence of increase in caregiver well-being	Minimal impact on caregiver well-being	Moderate impact on caregiver well-being	Significant impact on caregiver well-being	Strong evidence	Medium strength of evidence	Weak evidence	
		100								

Figure 2. The drug assessment framework (version 1.0). DFS indicates disease-free survival; HR, hazard ratio; OS, overall survival; PFS, progression-free survival; QOL, quality of life; RCT, randomized controlled trial.

TABLE 2. Summary Statistics of DAF Application

	Mean (Median)	SD (IQR)	Range
Total DAF score	94 (97)	39 (61)	18-179
Clinical benefit score	54 (52)	32 (46)	0-129

Abbreviations: DAF, drug assessment framework; IQR, interquartile ratio.

the Student *t* test for parametric data or the Wilcoxon test for nonparametric data. ESMO-MCBS scores had a nonparametric distribution and therefore Spearman correlation coefficients were used to assess the relation between ESMO-MCBS and DAF scores.

DAF scores were on a continuous scale and therefore the intraclass correlation coefficient (ICC) was used to assess inter-rater reliability for the DAF score, with an ICC value of 0 representing no agreement and a value of 1 representing perfect agreement. Weighted kappa was used to assess the inter-rater reliability for individual criterion scores because criterion scores were on a 4-level ordinal scale. For all analyses, the statistical significance level was set at .05. All statistical analyses were conducted by 2 authors (D.A.E. and L.L.) using SAS statistical software (version 9.4; SAS Institute Inc, Cary, North Carolina).

RESULTS

The final DAF was composed of 10 criteria: QOL, OS, unmet need, PFS, toxicity, RR, equity, feasibility, disease severity, and caregiver well-being (Fig. 2). The clinical benefit criteria were QOL, OS, PFS, toxicity, and RR. Weights totaled 100 and the highest weighted criterion was QOL (weighted at 19), followed by OS and unmet need (both weighted at 15). The clinical benefit criteria weights represented approximately 64% of the total weight. Total aggregate DAF scores can range from 0 to 300. Clinical benefit criteria scores can range from 0 to 192. When the DAF was applied to 60 drugs, the mean DAF score was 94 (range, 18-179) (Table 2).

Some data were missing for certain criteria. Fourteen of the drugs assessed did not have QOL data (Table 3)

TABLE 3. Application of the DAF Version 1.0, Drugs 1 to 30

Drug	Setting	QOL	OS	Unmet Need	PFS	Toxicity	RR	Equity	Feasibility	Disease Severity	Caregiver Well-Being	Total DAF Score	Clinical Benefit Score
1) Alectinib	First-line advanced NSCLC	0	NA	2	3	3	0	3	0	3	3	141	66
2) Inotuzumab ozogamicin	Acute lymphoblastic leukemia	0	0	1	3	0	2	0	0	3	NA	82	52
3) Trifluridine and tipiracil	Metastatic colorectal cancer	NA	1	2	0	0	0	0	0	2	1	59	15
4) Atezolizumab	Second-line advanced NSCLC	0	2	0	0	3	0	0	1	3	NA	81	60
5) Nivolumab	Classic Hodgkin lymphoma after failure of ASCT	2	NA	3	3	2	0	0	0	2	NA	149	94
6) Ribociclib	Advanced/metastatic breast cancer	0	0	1	3	0	1	0	3	2	NA	87	44
7) Olaratumab	Advanced soft-tissue sarcoma	NA	2	2	1	1	1	0	0	3	1	109	60
8) Regorafenib	Hepatocellular cancer	0	2	3	3	0	1	0	0	2	NA	129	74
9) Alectinib	Second-line advanced/metastatic NSCLC	0	NA	1	3	3	2	3	3	3	NA	148	82
10) Panitumumab	Left-sided metastatic colorectal cancer	0	0	1	1	0	NA	0	0	2	NA	37	12
11) Avelumab	Metastatic Merkel cell cancer	0	2	2	1	1	0	0	0	3	NA	97	52
12) Pembrolizumab	Metastatic urothelial cancer	2	1	3	0	2	0	0	0	2	NA	128	73
13) Venetoclax	Chronic lymphocytic leukemia	1	1	1	0	1	1	0	0	1	NA	72	52
14) Brentuximab vedotin	Hodgkin lymphoma recurrence risk or after ASCT	0	0	3	3	1	NA	0	0	2	NA	101	46
15) Fulvestrant	Advanced/metastatic breast cancer	0	1	0	2	1	0	0	0	2	NA	59	49
16) Pembrolizumab	Classic Hodgkin lymphoma	1	1	3	2	1	1	0	0	2	NA	131	76
17) Irinotecan liposome	Metastatic pancreatic cancer	0	2	2	3	0	1	0	0	3	NA	119	74
18) Midostaurin	Acute myeloid leukemia	NA	3	2	3	0	1	0	0	3	NA	134	89
19) Nivolumab and ipilimumab combination	Metastatic melanoma	0	2	2	2	0	1	0	0	3	NA	107	62
20) Dabrafenib and trametinib combination	Advanced/metastatic NSCLC	NA	1	1	1	0	0	0	3	3	2	83	27
21) Daratumumab	Multiple myeloma second-line and beyond	0	0	1	3	0	1	0	0	2	NA	69	44
22) Olaparib	Ovarian cancer	0	0	3	2	0	NA	0	2	3	NA	96	24
23) Blinatumomab	Adult acute lymphoblastic leukemia	1	3	1	3	0	1	0	0	3	NA	138	108
24) Nivolumab	Head and neck squamous cell carcinoma	0	2	1	0	3	0	0	0	2	NA	85	60
25) Rituximab	Acute lymphoblastic leukemia	NA	0	1	0	0	0	0	0	3	NA	30	0
26) Blinatumomab	Pediatric acute lymphoblastic leukemia	NA	1	3	0	1	0	1	0	3	1	95	25
27) Pembrolizumab	First-line advanced/metastatic NSCLC	0	2	2	3	3	1	0	0	3	NA	149	104
28) Ixazomib	Multiple myeloma	0	0	2	0	1	1	0	2	2	NA	70	18
29) Axitinib	Metastatic renal cell cancer	NA	0	1	0	1	NA	0	0	2	NA	35	10
30) Obinutuzumab	Follicular lymphoma	0	1	1	1	0	0	0	0	1	NA	47	27

Abbreviations: ASCT, autologous stem cell transplantation; DAF, drug assessment framework; NA, not available; NSCLC, non–small cell lung cancer; OS, overall survival; PFS, progression-free survival; QOL, quality of life; RR, response rate.

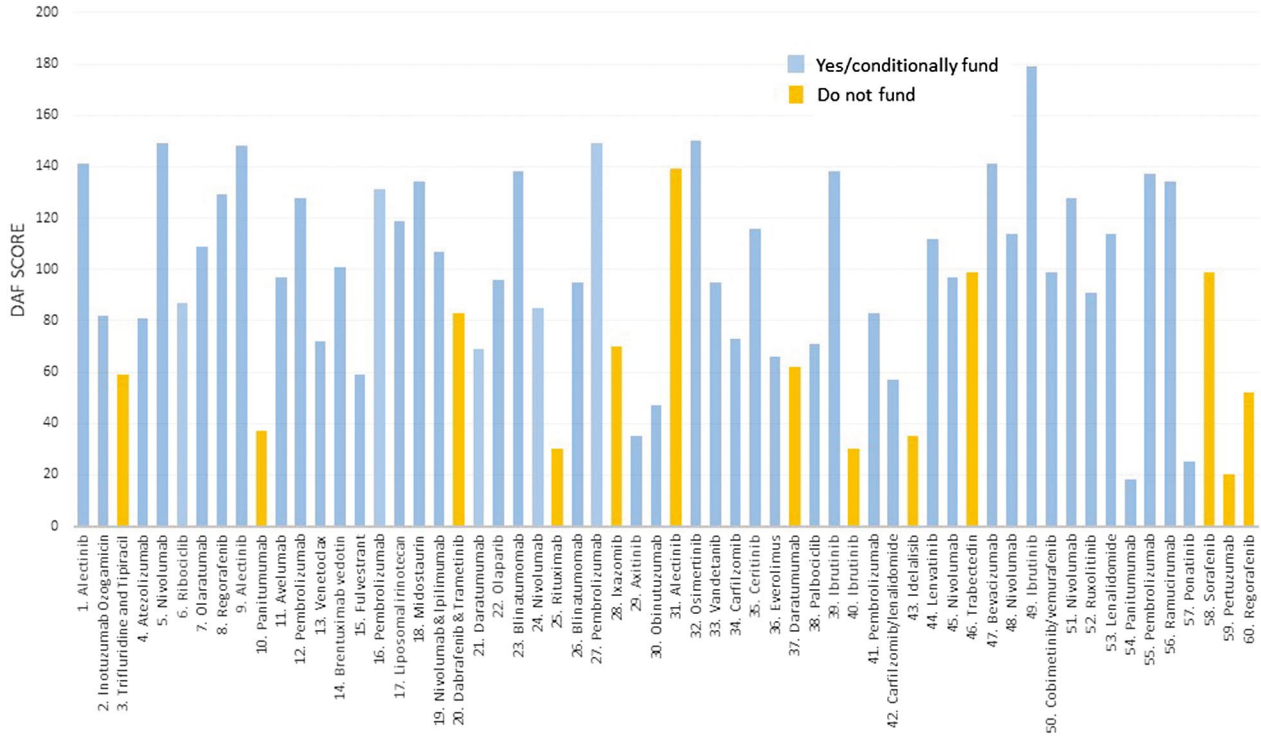

Figure 3. Drug assessment framework (DAF) scores and pan-Canadian Oncology Drug Review (pCODR) funding recommendations. Scores for drugs recommended for funding (yes/conditionally fund) are shown in blue and scores for drugs not recommended for funding (do not fund) are shown in orange (see Supporting Information Materials for the accompanying table with corresponding indications for each drug).

(see Supporting Information Materials for application of the remaining 30 drugs). Approximately 85% of the drugs did not report data regarding caregiver well-being. Other criteria with missing data were OS (7 drugs), RR (4 drugs), and PFS (1 drug).

Association Between DAF Scores and pCODR Funding Decision

Drugs fully or conditionally recommended for funding by pCODR had higher DAF scores than drugs not recommended (mean DAF score of 103 vs 63; Student *t* test $P = .0007$) (Fig. 3).

SOE modifier points were deducted for a variety of reasons, including noncomparative, nonrandomized trial design; post hoc analyses; and unconventional statistical analysis metrics (eg, high type I error rates). Thus, SOE point deductions did not apply to all criteria. SOE points were deducted only for the following criteria: QOL, OS, PFS, and RR. Drugs fully or conditionally recommended for funding tended to have fewer SOE points deducted compared with drugs not recommended for funding (median total SOE points deducted, 0 vs 24; Wilcoxon $P = .03$).

Correlation Between DAF and ESMO-MCBS Scores

The ESMO-MCBS (version 1.1) framework was applied to drugs for solid tumors that could generate a numerical score (37 drugs). The correlation coefficient between the total DAF and ESMO-MCBS scores was 0.26 (95% CI, -0.071 to 0.54; $P = .12$) (Fig. 4). The clinical benefit criteria scores were found to be mildly correlated with ESMO-MCBS scores (correlation coefficient, 0.33; 95% CI, 0.009-0.59 [$P = .045$]).

Sensitivity Analysis

When the subjective criteria scores were varied, the sensitivity analysis results supported the base case findings. Drugs with positive pCODR funding recommendations had significantly higher total DAF and clinical benefit scores. Mean DAF scores ranged from 64 (lower limit) to 143 (upper limit). In both the lower and upper limit sensitivity analyses, higher DAF scores were found

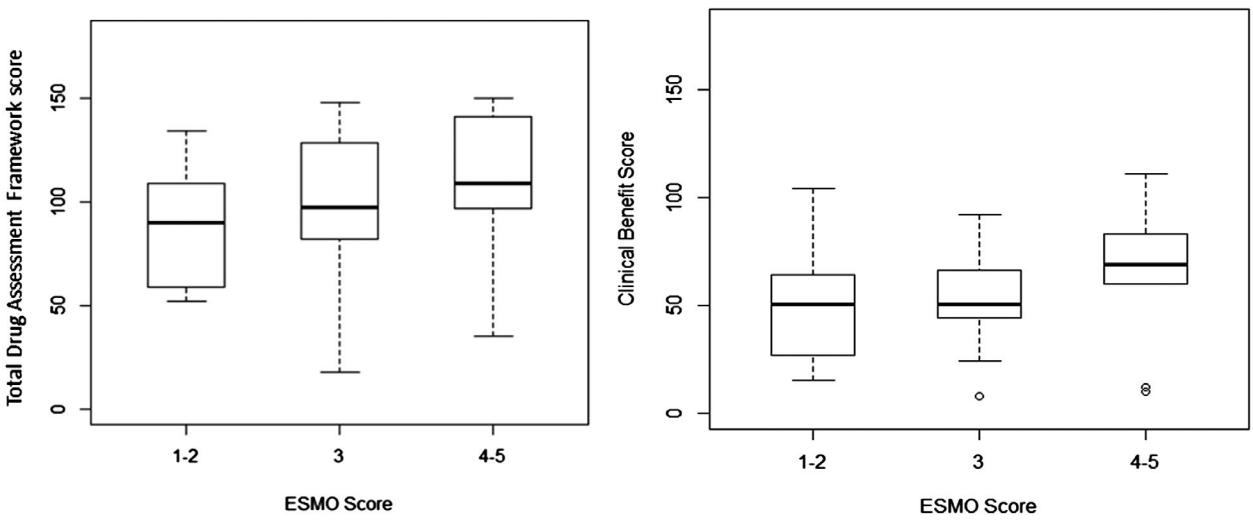

Figure 4. Correlation of the drug assessment framework and European Society for Medical Oncology Magnitude of Clinical Benefit Scale (ESMO-MCBS; version 1.1) scores.

TABLE 4. Weighted Kappa Scores for Individual Criterion Scores in the DAF

Criterion	Weighted Kappa[a]	95% CI
QOL	0.82	0.70-0.94
OS	0.96	0.91-1.0
Unmet clinical need	0.73	0.58-0.88
PFS	1.00	—
Toxicity	0.86	0.78-0.95
RR	0.83	0.67-0.96
Equity	1.00	—
Feasibility	0.90	0.77-1.00
Disease severity	0.80	0.69-0.92
Caregiver well-being	0.40[b]	0.05-0.75[b]

Abbreviations: DAF, drug assessment framework; OS, overall survival; PFS, progression-free survival; QOL, quality of life; RR, response rate.
[a]Weighted kappa scores were derived from the application of the DAF to 60 drugs.
[b]Values for caregiver well-being were low due to a large amount of missing data for this criterion (51 missing scores).

to be associated with drugs that were recommended for funding (lower limit DAF score: 71 vs 38 [Student *t* test $P = .0006$]; and upper limit DAF score: 152 vs 112 [Student *t* test $P < .0001$]).

Scenario analysis found that the mean clinical benefit score was 54 (range, 0-129) (Table 2). Drugs recommended for funding by pCODR had higher clinical benefit scores (mean clinical benefit score: 63 vs 23; Wilcoxon $P < .0001$).

Inter-Rater Reliability

Inter-rater reliability for DAF scores was found to be good (ICC, 0.93; 95% CI, 0.89-0.96). Weighted kappa demonstrated good agreement between scorers for each criterion (Table 4).

DISCUSSION

Rapidly escalating cancer drug prices present resource allocation issues that call for careful consideration of value for money. Assessment of incremental value has become increasingly important prior to the implementation and use of new cancer treatments. In the current study, MCDA was used to provide a comprehensive assessment of value as represented by a diverse group of stakeholders. The developed framework used 10 weighted criteria and assigned a score between 0 and 300, with higher scores representing high-impact drugs. Stakeholders weighted QOL, OS, and unmet need highest, followed by PFS, toxicity, RR, equity, feasibility, disease severity, and caregiver well-being. Higher DAF scores were found to be associated with positive pCODR funding recommendations, and clinical benefit scores were weakly correlated with ESMO-MCBS scores. Outside of the fundamentals of clinical benefit and cost, stakeholders identified additional criteria (unmet need, equity, feasibility, disease severity, and caregiver well-being) as important in cancer drug decision-making processes. Our developed framework assessed both curative-intent and palliative-intent therapies because stakeholders ultimately valued the magnitude of benefit more highly than the intent of therapy. Clinical trial data may not have fully captured long-term benefits appreciated from curative-intent therapy. Economic

analyses help to project magnitude of benefits into the future to better capture long-term benefits. Thus, the results of the final calculated DAF score can be used alongside the incremental cost, incremental cost-effectiveness threshold, and/or budget impact analysis to inform clinical decision making and funding decisions and establish explicit funding priorities. With the involvement of key perspectives in criteria selection and weighting, our developed framework is an important step in summarizing multiple perspectives of value and facilitating transparent decision making.

Both private and publicly funded health care systems can benefit from MCDA for health technology assessment and prioritization decisions.[18] Quantitatively ranking treatment options using an explicit process provides a summary measure that simplifies complex decisions. The DAF can be applied to a group of drugs submitted for funding to help identify highly ranked drugs that, within the setting of favorable economic profiles, then can be prioritized for funding. Because the development of the DAF was not initiated by pCODR, there are several potential areas for application of the DAF, including priority setting at the provincial level and informing pricing negotiations for drugs. Despite the increasing cancer drug budget over time, jurisdictions are challenged with funding all new treatments that receive a positive recommendation for funding. The DAF can help decision makers prioritize the highest impact (highest scoring) drugs. Although the calculated score summarizes overall impact and quality of evidence, the decision-making deliberations and discussions still are necessary.

Recently developed value frameworks were designed using different approaches, and some were for slightly different purposes.[7,8] Although the ESMO-MCBS was developed for similar public policy decision-making purposes as our framework, the magnitude of clinical benefit was the metric derived by the ESMO tool. Our clinical benefit scores only weakly correlated with ESMO-MCBS scores. Although some data suggested weak to moderate correlation between ASCO and ESMO framework scores, many are increasingly recognizing that value depends on the perspective of the assessor and the unique clinical context being considered.[29-32] Consequently, multiple perspectives of value are important in decisions regarding funding and sustainability. The ESMO-MCBS scale also places less importance on RR and PFS for treatments that demonstrate an improvement in OS. Data from the current study have demonstrated that, although OS was one of the most highly valued criteria, stakeholders still assigned value to PFS and RR even when a treatment demonstrated an OS benefit. We believe that this provides greater discriminant value to our DAF. With the MCDA process, multiple viewpoints are incorporated into a common framework that describes what is valued in a new treatment. As cancer treatment paradigms evolve, approaching value considerations from a perspective that allows patient, public, and clinician expectations to be incorporated can enhance the development of randomized clinical trials. Implementation of more tools such as the DAF that explicitly assess what matters most to public payers can better align drug developer and other stakeholder priorities to achieve value-based cancer care for patients.

Skedgel et al[12] reviewed pCODR decision summaries and found that, although criteria weights were implicit, the following factors carried the greatest weight in positive recommendations for pCODR funding: quality of the clinical evidence, relative survival gains, toxicity, and unmet need. It is interesting to note that these criteria fall into the 5 highest ranked criteria using our MCDA process. We implemented SOE as a variable that could modify any criterion. Field testing demonstrated that the SOE modification was relevant for only 4 of the clinical benefit criteria: QOL, OS, PFS, and RR. The combined influence of these 4 criteria on pCODR funding recommendations further indicates the impact of quality of evidence on funding recommendations. In recent years, an increasing number of drugs have been granted regulatory approval on the basis of data from studies lacking a comparator arm, typically within the setting of uncommon malignancies or molecular subtypes of certain cancers.[8] The SOE modifier provides a formal approach with which to assess single-arm and early-phase studies while acknowledging the limitations of the data. During the development of the framework, much discussion went into whether budget constraints should be considered by including cost as a criterion. The concern with a cost criterion is that it can obscure the assessment of value because a given DAF score would be the result of a mix of the benefits and the cost of a drug. Furthermore, determining the weight of a cost criterion would be challenging. Should the cost be considered as important as the sum of the benefits? Or should the cost be more or less important? Assigning a score to a cost criterion also might result in unnecessary loss of information. For example, a score of 0 may correspond to a drug costing >$200,000 per patient regardless of whether the drug costs $201,000 or $350,000. Cost is not fixed and remains subject to temporal and geographic variations. In prior studies using multicriteria decision analysis to assist health care decisions, only a minority used cost or cost-effectiveness as a criterion.[18] In addition to the challenges noted above, the

authors in these prior studies noted the risk of "double-counting" with a cost-effectiveness criterion because the effect dimension already is being captured by the clinical benefit criteria within the same framework. With these complexities noted, traditionally accepted measures of cost (ie, incremental cost-effectiveness threshold and budget impact analyses) were reserved to be compared with DAF scores in deliberative discussions.

Some limitations of the current study should be noted. The pCODR reports were descriptive and therefore interpretation of criteria such as unmet need, disease severity, equity, feasibility, quality of evidence, and caregiver well-being introduced some subjectivity. These risks were managed by using 3 independent reviewers, the resolution of disagreement by consensus, and standard qualitative coding techniques. In addition, uncertainty analysis, a fundamental component of MCDA, was performed. Base case and uncertainty analyses resulted in the same conclusions, which lends validity to the ultimate findings. Furthermore, the entire framework should be reviewed periodically to incorporate geographic and temporal changes in stakeholder principles and values.[33] Although our stakeholder group represented a wide spectrum of interests, it would be interesting to explore how criteria identification and weighting would differ by soliciting input from a larger population. Extension of this work can explore the impact of a larger stakeholder group by broadening the consultation to reach out to a larger population. Future work also can explore the impact of uncertainty by using different weighting techniques, stakeholder groups, and uncertainty analyses. Critics of MCDA note that logistical considerations, such as training decision makers to optimally use the tool, can lengthen the adoption process.[18] Due to individual cognitive biases, decision makers may have different ways of understanding the data and interpreting the scales. Expert training often is required to ensure the appropriate application of value assessment tools, and this issue has been recognized with the ASCO and ESMO value frameworks as well.[34,35]

In the current study, a criterion-based valuation framework was designed using multiple perspectives, and the robustness of the tool was demonstrated when compared with past submissions. This derived score represents the overall impact of a new cancer treatment, and the quality of evidence used to generate the score. This is the first version of the DAF, a dynamic tool that will be revised as public feedback is ascertained and the landscape of cancer treatment evolves. The DAF lends transparency and consistency to complex decision making. The DAF quantifies stakeholders' expectations of meaningful advances in patient outcomes, and is an important step toward the delivery of high-value cancer care.

FUNDING SUPPORT

Supported in part through a 2016 Canadian Association of Medical Oncologists' Fellowship Grant (to Doreen A. Ezeife), a 2017 Ontario Student Opportunity Trust Fund Grant (to Doreen A. Ezeife), a 2018 Canadian Institutes of Health Research Frederick Banting and Charles Best Master's Grant (to Doreen A. Ezeife), and in part by the Department of Medical Oncology of the Princess Margaret Cancer Centre.

CONFLICT OF INTEREST DISCLOSURES

Doreen A. Ezeife has received honoraria from Pfizer and Oncology Education for work performed outside of the current study. Don Husereau has received personal fees from the Canadian Agency for Drugs and Technology in Health (CADTH) for acting as a former member of a pan-Canadian oncology drug review expert review committee, personal fees from the Ontario Ministry of Health and Long-Term Care for acting as a former member of the Ontario Ministry Committee to Evaluate Drugs, and personal fees from AbbVie and Janssen for having acted as a paid advisor for work performed outside of the current study. Farzad Ali is an employee of and shareholder in Pfizer Inc; all opinions or statements conveyed in this publication are those of the authors and not necessarily those of Pfizer Inc. Peter Michael Ellis has received honoraria for acting as a paid member of the Speakers Bureau and Advisory Board for AstraZeneca, Pfizer, Bristol-Myers Squibb, and AbbVie and honoraria for acting as a paid member of the Advisory Board for Takeda for work performed outside of the current study. The other authors made no disclosures.

AUTHOR CONTRIBUTIONS

All authors were involved in the conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, software, supervision, validation, visualization, writing, review, and editing of the article.

REFERENCES

1. Prasad V, De Jesus K, Mailankody S. The high price of anticancer drugs: origins, implications, barriers, solutions. *Nat Rev Clin Oncol*. 2017;14:381-390. doi:10.1038/nrclinonc.2017.31
2. Tibau A, Molto C, Ocana A, et al. Magnitude of clinical benefit of cancer drugs approved by the US Food and Drug Administration. *J Natl Cancer Inst*. 2018;110:486-492. doi:10.1093/jnci/djx232
3. Ocana A, Tannock IF. When are positive clinical trials in oncology truly positive? *J Natl Cancer Inst*. 2011;103:16-20. doi:10.1093/jnci/djq463
4. Downing NS, Aminawung JA, Shah ND, Braunstein JB, Krumholz HM, Ross JS. Regulatory review of novel therapeutics–comparison of three regulatory agencies. *N Engl J Med*. 2012;366:2284-2293. doi:10.1056/nejmsa1200223
5. Ezeife DA, Truong TH, Heng DY, Bourque S, Welch SA, Tang PA. Comparison of oncology drug approval between Health Canada and the US Food and Drug Administration. *Cancer*. 2015;121:1688-1693. doi:10.1002/cncr.29246
6. Schnipper LE, Davidson NE, Wollins DS, et al; American Society of Clinical Oncology. American Society of Clinical Oncology Statement: A Conceptual Framework to Assess the Value of Cancer Treatment Options. *J Clin Oncol*. 2015;33:2563-2577. doi:10.1200/JCO.2015.61.6706
7. Schnipper LE, Davidson NE, Wollins DS, et al. Updating the American Society of Clinical Oncology Value Framework: revisions and reflections in response to comments received. *J Clin Oncol*. 2016;34:2925-2934. doi:10.1200/JCO.2016.68.2518
8. Cherny NI, Dafni U, Bogaerts J, et al. ESMO-Magnitude of Clinical Benefit Scale version 1.1. *Ann Oncol*. 2017;28:2340-2366. doi:10.1093/annonc/mdx310

9. Cherny NI, Sullivan R, Dafni U, et al. A standardised, generic, validated approach to stratify the magnitude of clinical benefit that can be anticipated from anti-cancer therapies: the European Society for Medical Oncology Magnitude of Clinical Benefit Scale (ESMO-MCBS). *Ann Oncol*. 2015;26:1547-1573. doi:10.1093/annonc/mdv249
10. Cheung MC, Sabharwal M, Chambers A, Han D, Sabarre KA, Chan K. Multiple dimensions of value: evaluative frameworks for new cancer therapies. *J Clin Oncol*. 2016;34:1428-1429. doi:10.1200/JCO.2015.66.4201
11. Doyal L, Tobias J. Informed consent in medical research. *BMJ*. 1998;316:1000-1005.
12. Skedgel C. The prioritization preferences of pan-Canadian oncology drug review members and the Canadian public: a stated-preferences comparison. *Curr Oncol*. 2016;23:322-328. doi:10.3747/co.23.3033
13. Ezeife DA, Kirk V, Chew DS, et al. Economic analysis of osimertinib in previously untreated EGFR-mutant advanced non–small cell lung cancer in Canada. *Lung Cancer*. 2018;125:1-7. doi:10.1016/j.lungcan.2018.08.024
14. Aguiar PN Jr, Haaland B, Park W, Tan PS, Del Giglio A, de Lima Lopes G Jr. Cost-effectiveness of osimertinib in the first-line treatment of patients with EGFR-mutated advanced non–small cell lung cancer. *JAMA Oncol*. 2018;4:1080-1084. doi:10.1001/jamaoncol.2018.1395
15. Marsh K, Ijzerman M, Thokala P, et al; ISPOR Task Force. 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;19:125-137. doi:10.1016/j.jval.2015.12.016
16. 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;19:1-13. doi:10.1016/j.jval.2015.12.003
17. Baran-Kooiker A, Czech M, Kooiker C. Multi-criteria decision analysis (MCDA) models in health technology assessment of orphan drugs–a systematic literature review. Next steps in methodology development? *Front Public Health*. 2018;6:287. doi:10.3389/fpubh.2018.00287
18. 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;32:345-365. doi:10.1007/s40273-014-0135-0
19. Angelis A, Montibeller G, Hochhauser D, Kanavos P. Multiple criteria decision analysis in the context of health technology assessment: a simulation exercise on metastatic colorectal cancer with multiple stakeholders in the English setting. *BMC Med Inform Decis Mak*. 2017;17:149. doi:10.1186/s12911-017-0524-3
20. Outcomes of cancer treatment for technology assessment and cancer treatment guidelines. American Society of Clinical Oncology. *J Clin Oncol*. 1996;14:671-679.
21. Browman GP, Manns B, Hagen N, Chambers CR, Simon A, Sinclair S. 6-STEPPPs: a modular tool to facilitate clinician participation in fair decisions for funding new cancer drugs. *J Oncol Pract*. 2008;4:2-7. doi:10.1200/JOP.0812001
22. Kwon SH, Park SK, Byun JH, Lee EK. Eliciting societal preferences of reimbursement decision criteria for anti cancer drugs in South Korea. *Expert Rev Pharmacoecon Outcomes Res*. 2017;17:411-419.
23. Leung L, de Lemos ML, Kovacic L. A deliberative framework to identify the need for real-life evidence building of new cancer drugs after interim funding decision. *J Oncol Pharm Pract*. 2017;24:584-598. doi:10.1177/1078155217722047
24. MacLeod TE, Harris AH, Mahal A. Stated and revealed preferences for funding new high-cost cancer drugs: a critical review of the evidence from patients, the public and payers. *Patient*. 2016;9:201-222.
25. Regier DA, Bentley C, Mitton C, et al. Public engagement in priority-setting: results from a pan-Canadian survey of decision-makers in cancer control. *Soc Sci Med*. 2014;122:130-139.
26. Zuniga MA, Carrillo-Zuniga G, Seol YH, Fos PJ. Multi-criteria assessment of county public health capability disparities. *J Health Hum Serv Admin*. 2009;32:238-258.
27. Sobrero AF, Pastorino A, Sargent DJ, Bruzzi P. Raising the bar for antineoplastic agents: how to choose threshold values for superiority trials in advanced solid tumors. *Clin Cancer Res*. 2015;21:1036-1044. doi:10.1158/1078-0432.CCR-14-1505
28. pCODR Drug Funding Recommendation. Accessed June 7, 2018. https://www.cadth.ca/sites/default/files/pcodr/pcodr_olaparib_lynparza_resub_fn_rec.pdf
29. Nabhan C, Phillips EG, Feinberg BA. Value in oncology: it is in the eyes of the beholder. *J Natl Compr Canc Netw*. 2019;17:2-5. doi:10.6004/jnccn.2018.7092
30. Evans WK, Cheung MC, Chan KK. Measuring value and benefit–a matter of perspective. *Lancet Oncol*. 2017;18:839-840. doi:10.1016/s1470-2045(17)30423-0
31. Cheng S, McDonald EJ, Cheung MC, et al. Do the American Society of Clinical Oncology Value Framework and the European Society of Medical Oncology Magnitude of Clinical Benefit Scale measure the same construct of clinical benefit? *J Clin Oncol*. 2017;35:2764-2771. doi:10.1200/JCO.2016.71.6894
32. Del Paggio JC, Sullivan R, Schrag D, et al. Delivery of meaningful cancer care: a retrospective cohort study assessing cost and benefit with the ASCO and ESMO frameworks. *Lancet Oncol*. 2017;18:887-894. doi:10.1016/S1470-2045(17)30415-1
33. Dionne F, Mitton C, Dempster B, Lynd LD. Developing a multi-criteria approach for drug reimbursement decision making: an initial step forward. *J Popul Ther Clin Pharmacol*. 2015;22:e68-e77.
34. Cherny NI, de Vries EGE, Schilsky RL. Letter to the Editor: when expertly applied, ESMO-MCBS and ASCO Net Health Benefit Scores usually agree. *J Natl Compr Canc Netw*. 2019;17:xxi. doi:10.6004/jnccn.2019.7293
35. Cherny NI, de Vries EGE, Dafni U, et al. Comparative assessment of clinical benefit using the ESMO-Magnitude of Clinical Benefit Scale Version 1.1 and the ASCO Value Framework Net Health Benefit Score. *J Clin Oncol*. 2019;37:336-349. doi:10.1200/jco.18.00729