Validity and Reliability of Three Value Frameworks for Oncology Therapeutics: A Pilot Study

Tanya G.K. Bentley, PhD,[1] Joshua T. Cohen, PhD,[2] Elena B. Elkin, PhD,[3] Julie Huynh, MD,[4] Arnab Mukherjea, DrPH, MPH,[5] Thanh H. Neville, MD, MSHS,[6] Ioana Popescu, MD, MPH,[6] Jenelle M. Zambrano DNP, CNS, RN,[1] Jackie Lee, BS,[1] Eunice Chang, PhD,[1] Michael S. Broder, MD, MSHS[1]

[1]Partnership for Health Analytic Research, LLC, Beverly Hills, CA, [2]Tufts Medical Center, Boston, MA, [3]Memorial Sloan-Kettering Cancer Center, New York, NY, [4]Harbor-UCLA Medical Center, Torrance, CA, [5]California State University, East Bay, Hayward, CA, [6]David Geffen School of Medicine at UCLA, Department of Medicine, Los Angeles, CA

BACKGROUND

- In response to rising spending in oncology care, various frameworks have been developed to assess the value of oncology drugs.
- These organizations include the American Society of Clinical Oncology (ASCO), European Society for Medical Oncology (ESMO), and Institute for Clinical and Economic Review (ICER).
- Despite their common goals, it is unclear whether the frameworks actually provide valid and reliable measurements of value and how to assess such validity and reliability in practice.

OBJECTIVE

- In this pilot study, we evaluated the validity and reliability of three value frameworks to understand the extent to which these tools can facilitate value-based treatment decisions in oncology.

Figure 1. Value Frameworks

Figure 1A. ASCO

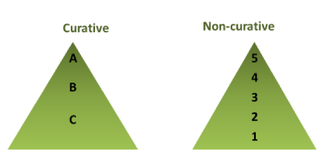
ASCO Framework for Assessing Value in Cancer Care

THE ASCO VALUE FRAMEWORK: ADVANCED DISEASE

Step 1: Determine the regimen's CLINICAL BENEFIT							
1.A. If Overall Survival (OS) reported?	YES. Assign an OS Score (1 through 5 as shown below) and multiply by 16. Write this number in the box labeled, "OS Score." **Proceed to 1.D.**						OS Score
	OS Score	1	2	3	4	5	
	Improvement in median OS (% change in median OS)	> 0%-24%	25%-49%	50%-75%	76%-100%	At double the median OS of new regimen, there is a 50% improvement in the fraction of patients surviving	
	NO. Proceed to 1.B.						
1.B. If OS is not reported, is Progression-Free Survival (PFS) reported?	YES. Assign a PFS Score (1 through 5 as shown below) and multiply by 11. Write this number in the box labeled, "PFS Score." **Proceed to 1.D.**						PFS Score
	PFS Score	1	2	3	4	5	
	Improvement in median PFS (% change in median PFS)	> 0%-24%	25%-49%	50%-75%	76%-100%	At double the median PFS of new regimen, there is a 50% improvement in the fraction of patients without progression or death	
	NO. Proceed to 1.C.						
1.C. If neither OS nor PFS is reported, is Response Rate (RR) reported?	YES. Assign an RR Score (1 through 5 as shown below) and multiply by 8. RR should be calculated by adding the complete response (CR) and partial response (PR) rates. Write this number in the box labeled, "RR Score." Proceed to 1.D.						RR Score
	RR Score	1	2	3	4	5	
	What was the reported response rate (CR + PR)?	> 0%-20%	21%-40%	41%-60%	61%-80%	81%-100%	
1.D. **Calculate the Clinical Benefit Score**	Insert the OS, PFS, or RR Score. **Note: You should have EITHER an OS Score OR a PFS score OR an RR score, NOT MORE THAN ONE.** Write the total in the box labeled "Clinical Benefit Score." The maximum allowable points are 80. **Proceed to Step 2.**						Clinical Benefit Score 0
Step 2: Determine the regimen's TOXICITY							
Calculate the Toxicity Score	For the regimens being assessed, compare the number of grade 3-5 toxicities (ie, calculate the sum of toxicities of grade 3-5 reported for each regimen) and assign a Toxicity Score (-20 through +20 as shown below). The score will be based on the difference in toxicity between the two regimens. Write this number in the box labeled, "Toxicity Score." The maximum allowable toxicity points are 20. **Proceed to Step 3.**						Toxicity Score
	Toxicity Score	-20	-10	0	+10	+20	
	Does the new regimen represent an improvement in toxicity over the standard of care/comparator?	Substantially less well tolerated (75%- 100% increase in the number of grade 3-5 toxicities reported for the new regimen.)	Less well tolerated (50%-74% increase in the number of grade 3-5 toxicities reported for the new regimen.)	Toxicity is the same (less than 49% increase and up to 49% fewer toxicities are reported for the new regimen.)	Better tolerated (50%-74% decrease in the number of grade 3-5 toxicities reported for the new regimen.)	Substantially better tolerated (75%-100% decrease in the number of grade 3-5 toxicities reported for the new regimen.)	
Step 3: Determine Bonus Points							
3.A. PALLIATION BONUS: Are data related to the palliation of symptoms reported?	YES. If a statistically significant improvement in cancer related symptoms is reported, award 10 points, and place this in the box labeled "Palliation Bonus Points." **Proceed to Step 3.B.**						Palliation Bonus Points
	NO. No bonus points are awarded. **Proceed to Step 3.B.**						
3. B. TREATMENT-FREE INTERVAL BONUS: Are data related to treatment-free interval reported?	YES. If a statistically significant improvement in treatment-free interval is reported, award points based on the table below, and place this in the box labeled "Clinical Benefit Bonus Points." This is the interval from completion of study treatment to initiation of next treatment. **Proceed to 3.C.**						Treatment-Free Interval Bonus
	Bonus Points	0	5	10	15	20	
	% Change	> 0%-19%	20%-35%	36%-49%	50%-74%	>75%	
	NO. No bonus points are awarded. **Proceed to Step 3.C.**						
3.C. Calculate Total Bonus Points	Add the Palliation Bonus Points (Step 3.A) and the Treatment-Free Interval Bonus Points (Step 3.B). Write this number in the box labeled "Total Bonus Points." The maximum points available for Bonus Points is 30. **Proceed to Step 4.**						Total Bonus Points 0
Step 4: Determine the regimen's NET HEALTH BENEFIT							
Calculate the Net Health Benefit	Add the Clinical Benefit Score (Step 1), Toxicity Score (Step 2), and Bonus Points (Step 3). This yields a Net Health Benefit Score. Write this number in the box labeled "Net Health Benefit." The maximum points available for Net Health Benefit are 130 (100 + 30 bonus points). Proceed to **Step 5.**						Net Health Benefit 0
Step 5: Determine the regimen's COST							
	Insert the drug acquisition cost (DAC) and patient co-pay based on how much the treatment regimen costs per month.						Cost Per Month: DAC: _____ Patient Co-Pay: _____
Step 6: Summary Assessment – Advanced Disease Framework							
Clinical Benefit	Toxicity	Bonus Points	Net Health Benefit	Cost (per month)			
0 /80	0 /20	0 /30	0 /130	DAC: $ 0.00 Patient Payment: _____			

ASCO, American Society of Clinical Oncology; CR, complete response; DAC, drug acquisition cost; OS, overall survival; PFS, progression-free survival; PR, partial response; RR, response rate.

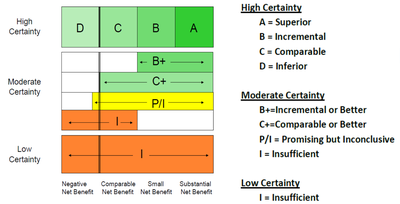
Figure 1B. ESMO



Curative - Evaluation form 1: for new approaches to adjuvant therapy or new potentially curative therapies

Non-curative -Evaluation forms 2a, b or c: for therapies that are not likely to be curative

Figure 1C. ICER



High Certainty
A = Superior
B = Incremental
C = Comparable
D = Inferior

Moderate Certainty
B+=Incremental or Better
C+=Comparable or Better
P/I = Promising but Inconclusive
I = Insufficient

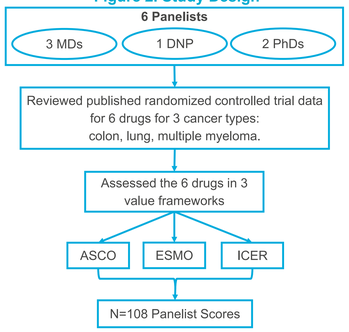
Low Certainty
I = Insufficient

METHODS

Value Framework Assessments

- We applied 3 frameworks (ASCO, ESMO, and ICER) to 6 drugs for 3 cancer types (colon, lung, and multiple myeloma).
  - 5 advanced cancer drugs
  - 1 adjuvant therapy drug
- Each assessment produced a single numeric or ordinal outcome (in aggregate the "panelist scores").
- Panelists were given a survey after completing the value assessments:
  - Rated different frameworks;
  - Provided comments regarding their experiences.

Figure 2. Study Design



Analyses

Validity

- Among the 5 advanced cancer drugs, we evaluated convergent validity: the correlation among drug rankings across frameworks.
  - Kendall's Coefficient of Concordance for Ranks (Kendall's *W*) was the statistical measure and used for the 5 advanced cancer drugs.
    1. Calculated mean scores for each drug.
    2. Ranked mean scores of each of the 5 drugs within each framework from highest to lowest.
    3. Compared rankings among the frameworks.
    4. Kendall's *W* ranges from 0 (no agreement) to 1 (complete agreement). P values tested alternative hypothesis of complete agreement ($W > 0$) against null hypothesis.
    5. Means were re-scaled to 0-100 for easy comparisons.

Reliability

- Inter-rater reliability measured the stability of frameworks' value estimates across users.
  - Intraclass correlation coefficients (ICC) with 95% confidence intervals (CI) were the statistical measure.
  - ICC was calculated separately for each framework.
  - ICC calculations were done assuming the 8 reviewers represent a random sample from a larger population of reviewers.

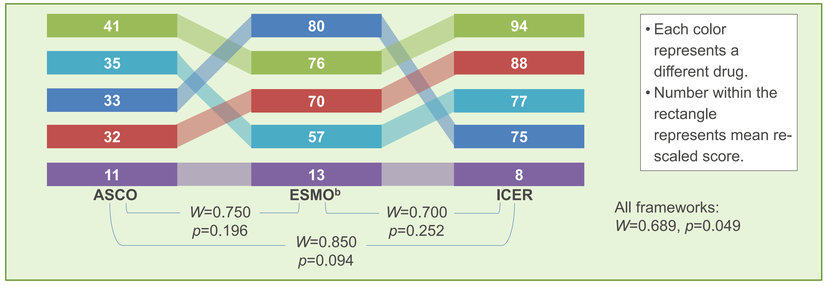
RESULTS

Overview

- Results are shown in Figure 3 (validity) and in the Table (reliability).
- Raw scores are on different scales and cannot be compared.
- When re-scaled from 0 (worst) to 100 (best), score ranges varied among frameworks.
- ICER had the widest range: 86 points, and ASCO had the narrowest range: 30 points.
  - ASCO: 11-41
  - ESMO: 13-80
  - ICER: 8-94
- ASCO scores were the lowest, and ICER scores were highest.
- Kendall's *W*=0.689

RESULTS (cont.)

Figure 3. Rankings and Re-Scaled Scores of 5 Cancer Drugs[a] Using 3 Frameworks



Columns represent each framework. Mean scores range from 0 to 100. Kendall's *W* is shown as a measure of concordance across all frameworks and each pairwise comparison.
[a] One of the drugs was not included in the rankings because it was an adjuvant therapy drug, and the rest were advanced cancer drugs; [b] An ESMO score of 0—outside of the standard ESMO range—was assumed when panelists had insufficient data to conduct the assessment.

Table. ICC (95% CI), All Reviewers

	ASCO	ESMO	ICER
All reviewers (n=6)	0.656 (0.945 - 0.316)	0.727 (0.960 - 0.407)	0.716 (0.958 - 0.370)

Panelists' Survey Results

- ESMO instructions were rated the clearest.
- ASCO was rated most logically organized.
- ESMO was rated to be the easiest to use.
- No single framework emerged as having the highest global panelist rating (e.g., comfort with using framework to assess treatment for a loved one).

CONCLUSIONS

- This is the first study to provide quantitative analyses of value assessment frameworks' validity and reliability.
- When applied to 6 oncology drugs:
  - Frameworks ranked similarly, indicating convergent validity (5 advanced cancer drugs only).
  - Overall, reliability was quite good.
- Our analysis ranked drugs across different cancers, although in practice, only within-cancer comparisons are useful. Future analyses will evaluate multiple drugs within each cancer type.
- The NCCN Evidence Blocks will also be included in future analyses.
- All frameworks should be refined using real-world testing and feedback, considering in particular the impact of using them to guide decisions on for patients.

Presented at SMDM 38th Annual North American Meeting, October 26, 2016, in Vancouver, BC, Canada