Verifying the value of existing frameworks for formulary review at a large academic health system: assessing inter-rater reliability

Brittany Lee Karas, PharmD; Mary Frances Picone, PharmD, BCPS; Shannon Werner, PharmD; and Megan Holsoopple, PharmD, BCPS

What is already known about this subject

- Value assessment frameworks (VAFs) are a way to measure and communicate the value of medications and other health care technologies for decision-making purposes.
- One critique of VAFs that is currently limiting their application to formulary decision making is whether they are able to provide consistent, reliable, and repeatable ratings among multiple reviewers.
- There are currently no studies globally assessing the inter-rater reliability for any medication reviewed for formulary or studies assessing inter-rater reliability for nononcology medications.

What this study adds

- The results suggest VAFs for the National Comprehensive Cancer Network, the American Society of Clinical Oncology, and the American College of Cardiology/American Heart Association are the most reliable and consistent evaluations.
- This is one of the first assessments of inter-rater reliability that applied 5 selected VAFs, including a larger number of medications than previous studies ($n=94$) and a large proportion of nononcology medications (68%).

Author affiliations

Brittany Lee Karas, PharmD, Department of Pharmacy, Northwestern Memorial Hospital, Chicago, IL. Mary Frances Picone, PharmD, BCPS, and Shannon Werner, PharmD, Center for Medication Utilization, Froedtert & the Medical College of Wisconsin, Milwaukee. Megan Holsoopple, PharmD, BCPS, RxRevu, Denver, CO.

AUTHOR CORRESPONDENCE:
Brittany Lee Karas, 312.695.4631;
Brittany.lee@nm.org

J Manag Care Spec Pharm.
2021;27(4):488-96

Copyright © 2021, Academy of Managed
Care Pharmacy. All rights reserved.

ABSTRACT

BACKGROUND: The value assessment framework (VAF) is one approach to assessing the evidence and value of medications. VAFs are a way to measure and communicate the value of medications and other health care technologies for decision-making purposes. Given the increasing number of high-cost medications, challenging formulary inquiries, and critiques of currently available tools, health systems need to explore a standardized way to incorporate value assessment into formulary decision making.

OBJECTIVES: To (a) evaluate existing VAFs by measuring inter-rater reliability among

typical clinicians completing formulary reviews and (b) explore general implications of applying these tools to formulary decision making for all medications at a large academic health system.

METHODS: This was a retrospective, observational study at a single health system. A list of medications added, denied, and removed from the system formulary from September 1, 2013, through August 31, 2018, was collected. Published VAFs, such as the American Society of Clinical Oncology (ASCO) Value Framework, European Society of Medical Oncology (ESMO) Magnitude of Clinical Benefit Scale, National Comprehensive Cancer Network (NCCN) Evidence Blocks,

American College of Cardiology/American Heart Association Value Framework, and the incremental cost-effectiveness ratio (ICER) calculation were applied by 3 different reviewer groups. The primary outcome was inter-rater reliability among the 3 different reviewers for a given framework. Cohen's weighted kappa and the intraclass correlation coefficient (ICC) were used to assess inter-rater reliability.

RESULTS: The frameworks were applied to 94 medications. The VAFs with the highest ICCs between all 3 raters were NCCN (0.635; 95% CI=0.387-0.823) and ASCO (0.634; 95% CI=0.370-0.832), both indicating moderate inter-rater reliability. The VAFs with the

lowest ICCs were ESMO (0.368; 95% CI=0.126-0.611) and ICER (0.159; 95% CI =0.018-0.365), with ICCs corresponding to poor reliability.

CONCLUSIONS: Because high-cost medications are a challenge to health systems, VAFs may be beneficial to target formulary decision making in this setting. Applying VAFs proactively may improve inter-rater reliability and usability in formulary decision making.

High-cost medication therapies have become a challenge for health care in America.^{1,2} In 2018, prescription drugs accounted for up to 17% of total health care costs in the United States, up from 15.3% in 2013.³⁻⁵ Because of these rising costs, different initiatives have been developed to discern the value of medications. The definition of “value” varies depending on the area of health care and the stakeholders involved. The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) recently used a pharmacoeconomic approach to define “value”: what one is willing to pay to acquire additional health care or services compared with the opportunity cost of what one is willing to give up in order to obtain the additional health care or services.⁶ Pharmacoeconomic studies currently explore the economic value of treatment options, but there is a need in the United States for a way to create and distribute improved information on the clinical and economic value of medications.⁷

Maintaining a formulary and creating a process for selecting medications for use within a health system is required by the Joint Commission and the National Integrated Accreditation for Healthcare Organizations for accredited hospitals.^{8,9} There are minimum criteria for pharmacy & therapeutics (P&T) committees, with the key principles being unbiased review of safety and efficacy.^{10,11} Although cost remains a necessary element, the concept of value may need consideration, given the rising cost of certain medications and subsequent cost to patients.

One approach to assess the evidence and value of medications is through the use of value assessment frameworks (VAFs).⁶ Different organizations, such as the American Society of Clinical Oncology (ASCO), European Society of Medical Oncology (ESMO), and National Comprehensive Cancer Network (NCCN), have created VAFs to aid patients, prescribers, and payers in determining the value of medications.¹²⁻¹⁴ These frameworks are a way to measure and communicate the value of medications and other health care technologies for decision-making purposes.

Current frameworks attempt to account for a variety of factors thought to affect the value of a medication, including the quality of clinical data, magnitude of treatment

effects, possibility of severe adverse events, ancillary benefits, and cost-effectiveness.¹⁵ The frameworks vary in scope and target audience, with a majority targeted towards patients and providers making decisions on an individual's treatment (e.g., ASCO) or payers and policy-makers determining the coverage of treatments at a larger level (e.g., ESMO).¹⁶ There is limited information available on the utility of VAFs for formulary decision making within hospitals or health systems.

With the cost of cancer care predicted to increase,¹⁷ institutions have started to explore ways to incorporate value assessment into formulary decision making for oncology agents with various published VAFs.^{18,19} One institution evaluated existing frameworks and decided to use 3 frameworks: the ASCO Value Framework, NCCN Evidence Blocks, and Institute for Clinical and Economic Review. These frameworks were applied to a pilot therapy for immunotherapy agents in previously treated advanced non-small cell lung cancer (NCSLC) and applied to multiple programmed death (PD-1) and programmed death-ligand 1 (PD-L1) inhibitors, including atezolizumab (Tecentriq), nivolumab (Opdivo), and pembrolizumab (Keytruda).

After applying these frameworks, atezolizumab was deemed the preferred agent for NCSLC with negative PD-L1 expression, since atezolizumab and nivolumab had similar efficacy and safety profiles, but atezolizumab was more cost-effective at the time of the pilot. Using the results from this pilot, the institution created a workflow using these 3 VAFs to assist in determining the value of select oncology agents. Since the pilot, nivolumab was approved for a flat dose of 480 mg every 4 weeks compared with a weight-based dose every 2 weeks; this approval may have altered the median cost and incremental cost estimates.¹⁸ Another institution developed a specialty drug subcommittee of the P&T committee to apply VAFs, such as NCCN Evidence Blocks and Institute for Clinical and Economic Review, along with pharmacoeconomic modeling, when reviewing formulary additions and requests for nonformulary drugs.¹⁹

One critique of VAFs currently limiting their application to formulary decision making is whether they are able to provide consistent, reliable, and repeatable ratings among multiple reviewers. Available literature on the inter-rater reliability of VAFs demonstrates conflicting results.²⁰⁻²² Bentley et al. (2017) measured the reliability of 4 VAFs: ASCO, ESMO, NCCN, and Institute for Clinical and Economic Review.²⁰ Eight panelists assessed 15 medications for advanced breast and lung cancers and castration-refractory prostate cancer using the 4 VAFs. Inter-rater reliability was assessed using intraclass correlation coefficients (ICCs). ASCO (ICC=0.800, 95% CI=0.660-0.913) and ESMO (ICC=0.818, 95% CI=0.686-0.921) demonstrated

excellent inter-rater reliability, while NCCN exhibited poor inter-rater reliability (ICC=0.153, 95% CI=0.045-0.371). Another study assessed the inter-rater reliability of the ASCO VAF for 11 oncology agents among 8 clinicians and found the inter-rater reliability to be slight.²² There are currently no studies globally that have assessed the inter-rater reliability for any medication reviewed for formulary or studies assessing inter-rater reliability for nononcology medications.

Given the increasing number of high-cost medications, challenging formulary inquiries, and critiques of the currently available tools, P&T committees need to explore a standardized way to incorporate value assessment into formulary decision making. There is some literature regarding inter-rater reliability in the setting of formulary decision making using a variety of oncology-focused VAFs, but literature is lacking for all-encompassing VAFs. The goals of this study were to (a) evaluate existing VAFs to measure inter-rater reliability among typical clinicians completing formulary reviews and (b) explore general implications in applying these tools to formulary decision making for all medications.

Methods

This was a retrospective, observational study at a single health system, composed of one 694-bed academic medical center and 2 community hospitals with a cumulative 272 beds. The academic medical center is 340B eligible, while the 2 community hospitals are not. Formulary decisions are determined for the health system at monthly System P&T Committee meetings, which evaluates clinical efficacy and safety, and budget impact is estimated based on medication cost and anticipated use. A team of 4 drug information pharmacists supports System P&T and serves to promote effective, financially responsible, and safe use of medications. This study was determined to be an internal institutional quality improvement project and not human subjects research; therefore, it did not need institutional review board review and approval.

A list of medications added, denied, and removed from the system formulary September 1, 2013–August 31, 2018, was collected using previous P&T committee meeting agendas and minutes. Medications were classified as oncology or nononcology based on the initial requested indication. Excluded items included biosimilars, medical devices, and formulary line-item (e.g., dosage forms and strengths) additions, modifications, and removals for existing formulary medications. The monographs (i.e., formulary reviews) for these medications were collected. Monographs at this institution are completed by the drug information pharmacists

or a learner (e.g., fourth-year pharmacy student or pharmacy resident). The monographs evaluate clinical efficacy and safety using comparator studies, placebo-controlled trials, and meta-analyses.

Budget impact is also assessed using the institution acquisition cost and estimated annual use. The monograph is completed before the review of the medication with the P&T committee. From the monograph, 1 piece of primary literature was selected with the highest level of evidence and with an effort to select the phase 3 trial or pivotal trial used for drug approval for the requested indication.

Published VAFs were reviewed by the study team for inclusion and selected based on usability and scope of application. Frameworks were excluded if the scope of the framework was very specific (e.g., only included certain medications) or was determined difficult to use (e.g., relied on patient-specific inputs). The VAFs included were the ASCO Value Framework, ESMO Magnitude of Clinical Benefit Scale, NCCN Evidence Blocks, and the American College of Cardiology/American Heart Association (ACC/AHA) Value Framework. Memorial Sloan Kettering Cancer Center's Drug Abacus and Institute for Clinical and Economic Review VAFs were excluded. In addition to the 4 selected VAFs, the incremental cost-effectiveness ratio (ICER) calculation was applied.

The 3 oncology-specific VAFs were ASCO, ESMO, and NCCN. The ASCO framework produces a net health benefit (NHB) score relative to a comparator (i.e., active therapy or placebo), incorporating clinical efficacy, toxicity, and bonus points determined by long-term survival, palliation, quality of life, and treatment-free interval.¹² Medications evaluated in a single-arm study were excluded from this framework. ESMO yields a result of grades 1 through 5, with 5 correlating with the most favorable grade of clinical efficacy, quality of life, and grade 3 and 4 toxicities.¹³ The NCCN framework ranks 5 categories, including efficacy, safety, quality of evidence, consistency of evidence, and affordability, from 1 to 5, with 5 being the highest grade.¹⁴ To create a single output for the NCCN framework, the outputs for each category were averaged to yield a single number as done in a previous study.²⁰

The remaining ACC/AHA framework and the ICER calculation were applied to all medications. The ICER calculation estimates cost using monetary units.²³ To determine the ICER for a specific medication, the difference in cost between the medication and a comparator is divided by the difference in outcome between the medication and a comparator. The comparator therapy and outcome was determined based on the piece of primary literature previously selected. Medications evaluated in a single-arm study were excluded from this framework. Costs for

FIGURE 1 Process for Determining Value Assessment Framework Application for Evaluated Medications

^aComparator studies included placebo-controlled trials and active-controlled trials.

ACC/AHA=American College of Cardiology/American Heart Association; ASCO=American Society of Clinical Oncology; ESMO=European Society of Medical Oncology; ICER=incremental cost-effectiveness ratio; NCCN=National Comprehensive Cancer Network; VAF=value assessment framework.

the medication and the comparator were the institution acquisition cost collected from the previously identified monograph and, thus, represent the costs expected at the time of review by System P&T. If the cost of the comparator was not listed in the monograph, the medication was excluded from this framework.

The ACC/AHA framework yields 2 results, the level for certainty of treatment effect (A, B, and C) and the class for size of treatment effect (I, IIa, IIb, III); this scoring system is used currently in ACC/AHA guidelines, and it has also been used most frequently by this institution in completing formulary reviews. ACC/AHA also proposes a level of value by incorporating the ICER equation and quality-adjusted life-years (QALYs) to determine value.²⁴ Because of the retrospective nature of this study, it was determined out of scope to review pharmacoeconomic literature for QALYs for all medications, since this information was unlikely to have been consistently part of a prepared monograph presented

to System P&T. Therefore, only the level and class were evaluated for ACC/AHA.

Included framework tools were applied by 3 different reviewer groups: a fourth-year pharmacy student, a drug information resident, and a drug information pharmacist. Since there were multiple fourth-year pharmacy students and drug information pharmacists available, medications were randomized to members of these groups using a random number generator and stratified by oncology and nononcology agents. Reviewers were given the published information on each framework and a flow diagram (Figure 1) to determine whether the medication should be excluded from a certain VAF. Figure 1 was also used by each reviewer to determine which framework to apply for a specific medication. To gain a better understanding of the VAFs, an initial training session was completed. All the reviewers independently completed each VAF with the

FIGURE 2 Medications Included for Application of Value Assessment Frameworks

^aExcluded medications included no published studies in the monograph (n=10), removals with no published literature (n=8), missing monograph (n=5), biosimilar (n=1), formulary line item (n=1), investigational biologic product (n=1), and medical device (n=1).

same oncology medication not included in this study. After completing the framework, the reviewers met as a group to discuss the VAFs and their findings. When reviewing differences in the VAF outputs, specific questions were addressed to develop a consistent approach, such as the interpretation of ASCO toxicities and the cost used to calculate the ICER.

The primary outcome was inter-rater reliability among the 3 different reviewers for a given framework. Tests of agreement between pairs of reviewers were performed using Cohen's weighted kappa. The ICC was used to assess agreement among all 3 raters. One pharmacist collated the information from the various different reviewers. The ICCs and their 95% CIs were calculated using a single-measure, consistency, two-way random effects model. All statistical analyses were performed using R, version 3.4.4 (R Foundation for Statistical Computing, Vienna, Austria).²⁵

There are no standard values to determine acceptable reliability, and a variety of reliability definitions exist.^{26,27} The following ICC ranges were used to assess reliability: <0.5 indicated poor reliability, 0.5–0.75 indicated moderate reliability, 0.75–0.9 indicated good reliability, and >0.9 indicated excellent reliability.²⁷ For the ASCO VAF, in addition to the NHB, ICCs were determined separately for the clinical

benefit score, which assesses the clinical efficacy and the toxicity score.

Results

A total of 121 medications were added, denied, or removed from the formulary from September 1, 2013, through August 31, 2018. Of the medications reviewed, 27 medications were excluded, with the majority of exclusions from this analysis due to absence of primary literature in the monographs or formulary removals with no published literature cited. Formulary removals generally did not require the completion of a monograph with published literature, since most removals were due to low utilization or replacing with a different formulary agent. Figure 2 illustrates the total number of additions, denials, and removals, as well as the number of oncology and nononcology medications. The frameworks were applied to 94 medications, including 30 oncology medications. A total of 98 evaluations were completed, with 4 medications (i.e., adalimumab [Humira], ceftazidime/avibactam [Avycaz], omalizumab [Xolair], and vedolizumab [Entyvio]) reviewed for 2 separate indications.

The ICCs and 95% CIs for 3 raters and the weighted Cohen's kappa between each rater are reported in Table 1.

TABLE 1 Weighted Cohen's Kappa and Intraclass Correlation Coefficients (95%) for Applied Value Assessment Frameworks

Framework (Number of Medications Evaluated)	ASCO (n=17)	ESMO (n=26)	NCCN (n=19)	ICER (n=43)	ACC/AHA (n=96 ^a)
Weighted Cohen's kappa for Rater 1 ^b & Rater 2 ^c	0.301	0.510	0.970	0.336	0.509
Weighted Cohen's kappa for Rater 1 ^b & Rater 3 ^d	0.427	0.211	0.427	0.197	0.290
Weighted Cohen's kappa for Rater 2 ^c & Rater 3 ^d	0.223	0.248	0.444	0.219	0.544
ICC for 3 raters (95% CI)	0.634 (0.370-0.832)	0.368 (0.126-0.611)	0.635 (0.387-0.823)	0.159 (-0.018-0.365)	0.515 (0.398-0.625)

^aOf the 98 reviews, 1 rater did not fill out the ACC/AHA for 2 medications.

^bRater 1: Drug information resident.

^cRater 2: Drug information pharmacist.

^dRater 3: Pharmacy student.

ACC/AHA=American College of Cardiology/American Heart Association; ASCO=American Society of Clinical Oncology; ESMO=European Society of Medical Oncology; ICC=intraclass correlation coefficient; ICER=incremental cost-effectiveness ratio; NCCN=National Comprehensive Cancer Network.

The VAFs with the highest ICCs between all 3 raters were NCCN (0.635; 95% CI=0.387-0.823) and ASCO (0.634; 95% CI=0.370-0.832), both indicating moderate inter-rater reliability. Following ASCO and NCCN, the ACC/AHA VAF was also considered to have moderate inter-rater reliability (0.515; 95% CI=0.398-0.625). The lowest ICCs were ESMO VAF (0.368; 95% CI=0.126-0.611) and the ICER calculation (0.159; 95% CI=-0.018-0.365), with ICCs corresponding to poor reliability. When reviewing weighted Cohen's kappa between pairs of raters, there was higher agreement among the drug information resident and the drug information pharmacist for the ESMO VAF and ICER calculation than either of these raters versus the pharmacy student (Table 1).

To further assess the ASCO VAF, the ICCs and weighted Cohen's kappa for the clinical benefit and toxicity components were calculated (Table 2). The ICC for the clinical benefit score was higher than the final NHB component for ASCO. The ICC for the toxicity score was a negative number, reflecting the true ICC to be very low.²⁸

Discussion

This is one of the first assessments of inter-rater reliability when applying the ASCO, ESMO, NCCN, and ACC/AHA VAFs, along with the ICER calculation, to medications reviewed for formulary at a large academic medical center. Before incorporating VAFs into formulary review at this institution, it was necessary to determine the reliability and usability of the VAFs. This assessment included a larger number of medications than previous studies, although included less reviewers.^{22,29} Nononcology medications represented the majority of included medications (68%; 64 of 94). The consensus among reviewers was NCCN, ESMO, and ACC/AHA

VAFs were easier to navigate and understand compared with the ASCO VAF and ICER calculation. This is relatively consistent with a previous study where reviewers were asked to rate their experiences with VAFs.²⁰

The NCCN, ASCO, and ACC/AHA frameworks were considered to have moderate inter-rater reliability, while ESMO and ICER were considered poor. Another study on inter-rater reliability with VAFs used different ICC ranges to assess reliability, where <0.40 represented poor reliability, 0.40-0.59 fair reliability, 0.60-0.74 good reliability, and ≥0.75 excellent reliability.²⁰ Using these definitions, NCCN and ASCO VAFs would be considered to have good inter-rater reliability. The higher reliability of NCCN and ASCO VAFs support the proposal by Saunders et al. (2019) to incorporate these 2 VAFs when reviewing oncology medications for formulary.¹⁸

Bentley et al. assessed the inter-rater reliability of ASCO, ESMO, and NCCN in 15 oncology medications with 8 reviewers.²⁰ The reviewers were provided with phase 3 randomized controlled trials; no single-arm studies were included. They concluded that ESMO and ASCO were the most reliable, with ICCs of ≥0.75 representing excellent reliability. The NCCN VAF in their study was found to have poor reliability (0.153; 95% CI=0.045-0.371). The study authors indicated that the poor reliability of NCCN was likely due to lack of instructions and lower discriminatory ability. For our study, reviewers were instructed to find the NCCN Evidence Block for the reviewed medication most closely matching the disease state, and inclusion criteria of the selected literature used for the other VAFs. With more direction and instruction on how to use the NCCN VAF, the reliability of the framework drastically improved.

TABLE 2 **Weighted Cohen's Kappa and Intraclass Correlation Coefficients (95%) for Component Scores of ASCO Value Assessment Framework**

	ASCO Clinical Benefit Score	ASCO Toxicity Score	ASCO Net Health Benefit
Weighted Cohen's kappa for rater 1 ^a & rater 2 ^b	0.819	0.490	0.301
Weighted Cohen's kappa for rater 1 ^a & rater 3 ^c	0.562	0.127	0.427
Weighted Cohen's kappa for rater 2 ^b & rater 3 ^c	0.498	0.066	0.223
ICC for 3 raters (95% CI)	0.822 (0.652-0.925)	-0.180 (-0.338-0.114)	0.634 (0.370-0.832)

^aRater 1: Drug information resident

^bRater 2: Drug information pharmacist

^cRater 3: Pharmacy student.

ASCO = American Society of Clinical Oncology; ICC = intraclass correlation coefficient.

In this study, the ESMO VAF was applied to all oncology medications, even if the selected literature to apply the VAF was a single-arm or noncomparator study. Some would argue that the ESMO VAF is only designed for use with comparator studies,²⁹⁻³² which could explain why the ESMO VAF was less reliable in this study than in previously reported studies, when only comparator studies were included.^{20,21} In 2017, ESMO included a new Form 3 for single-arm studies in orphan diseases or a disease with a high unmet need.¹³ With Form 3, the inclusion of noncomparator literature for our study was appropriate but may have introduced more variability than previously reported studies. Not all of the noncomparator studies assessed may have been appropriate for Form 3 and may have introduced bias when determining inter-rater reliability.

Consistent with other studies, the ICC for the ASCO clinical benefit score was greater than the ICC for toxicity.^{20,22} The instructions in the toxicity component for the ASCO VAF were found to be complicated and confusing, consistent with previous studies.^{12,22} When reviewing the ASCO VAF, the calculation of the percent difference between the toxicity score of the medication and the comparator was not the typical calculation for percent difference.¹² During training for reviewers on the ASCO VAF, this point was discussed, and it was determined to use the calculation in the published framework. Even with this discussion, there was low reliability in the toxicity portion.

Also, some components of the toxicity score were not consistently reported in studies. For example, persistent toxicity at 1 year was not always reported or provided in a table of reported adverse events, which contributed to variability in reviewer interpretation and decreased inter-rater reliability. The ASCO VAF would likely have higher reliability if the instructions for the toxicity portion were more detailed and clear, addressing how to handle unreported data.

LIMITATIONS

This study is not without limitations. The retrospective nature may have led to some of the variability and lack of consistency among the frameworks and the ICER calculation, which was significantly limited in what pricing information was reported in the monograph (e.g., different costs associated with different settings of care) and how costs or budget impacts were calculated (e.g., cost per dose, cost per unit, cost per year). Using these values to retrospectively calculate the cost per year or per treatment course likely affected the variability seen with this tool. Using the ICER calculation in real time could improve inter-rater reliability. In addition, there are no standard values to determine acceptable reliability and a variety of reliability definitions exist.^{26,27} One set of ICC ranges was selected to assess reliability in this study,²⁷ yet other ICC ranges could have been used and might have altered the conclusions.

Furthermore, the primary outcome reported in selected studies was not always overall survival or progression-free survival, rendering the ICER calculation less useful. It was challenging to interpret the clinical meaning of the ICER calculation if the primary outcome was the change in baseline for a certain measure, the proportion of patients achieving a goal, or a composite endpoint. The outcome life-years or QALYs allows for a meaningful output from the ICER calculation, since there are more established value cutoffs.^{33,34}

Another limitation that could have affected the results of this study is the varying level of baseline familiarity among reviewers. Although a brief training session and resources for the frameworks were provided, different interpretations of the frameworks likely existed; this may have contributed to lower ICCs than previously mentioned studies. The training process allowed reviewers to have similar experiences with the VAFs; however, clinical experiences of the reviewers were not standardized, especially for the fourth-year

pharmacy students. Weighted Cohen's kappa between raters (Table 1) support lower ICC and less reliability for fourth-year pharmacy students, which could be attributed to less clinical experience. Despite this being a study limitation, including fourth-year pharmacy students in the review was essential, since they often create monographs and therefore may use VAFs for future formulary reviews.

When considering the audience for these VAFs, it is important to consider that not all of these VAFs are targeted for policy decision making. Target stakeholders for the ESMO VAF and ICER calculation are payers and policymakers, while ASCO and NCCN are for patients and physicians.^{29,32,35,36} Although formulary decision making aligns more closely with policy decision making, other studies involving formulary decisions in academic medical centers have used ASCO or NCCN VAFs.^{18,19} Prospectively applying the VAFs to new medications requested for formulary addition would allow for feedback and discussion from the P&T committee members and further define the utility of VAFs. When incorporating VAFs into monographs and formulary decisions, P&T committee members will require education on the concept of VAFs, their limitations, and intended audience to ensure appropriate interpretation of the outputs. In addition, although VAFs can be a tool to assess value for formulary decisions, they should not be the sole input for decisions.³⁷

Conclusions

Innovative approaches are needed to tackle the challenges that high-cost medication therapies present to hospitals and health systems, patients, and providers throughout the United States. VAFs present a unique approach to assessing the value of medications. The most reliable VAFs

in the present study were ASCO, NCCN, and ACC-AHA with moderate reliability. Applying the frameworks prospectively may improve inter-rater reliability and the usability of VAFs to assist in formulary decision making. New VAFs or revisions to current VAFs may be beneficial to target formulary decision making in this setting. Future studies should focus on the correlation between VAFs and formulary decisions.

DISCLOSURES

No outside funding supported this study. The authors have nothing to disclose.

ACKNOWLEDGMENTS

Kasey Schmidt, PharmD, and Youqi Zhang, PharmD, contributed to reviewing the value assessment frameworks, and Lisa Rein provided statistical support.

REFERENCES

1. Kesselheim AS, Avorn J, Sarpatwari A. The high cost of prescription drugs in the United States: origins and prospects for reform. *JAMA*. 2016;316(8):858-71.
2. Neumann PJ, Cohen JT. Measuring the value of prescription drugs. *N Eng J Med*. 2015;373(27):2595-97.
3. Centers for Medicare & Medicaid Services, Office of the Actuary. National health care spending in 2016. Accessed March 7, 2021. <https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/Downloads/NHE-Presentation-Slides.pdf>
4. U.S. Department of Health & Human Services. Observations on trends in prescription drug spending. ASPE Issue Brief. March 8, 2016. Accessed March 7, 2021. <https://aspe.hhs.gov/system/files/pdf/187586/Drugspending.pdf>
5. Organisation for Economic Co-operation and Development. Pharmaceutical spending (indicator). Retrieved March 29, 2019. Accessed March 7, 2021. <https://data.oecd.org/healthres/pharmaceutical-spending.htm>

6. Garrison LP Jr, Neumann PJ, Willke RJ, et al. A health economics approach to U.S. value assessment frameworks-summary and recommendations of the ISPOR special task force report [7]. *Value Health*. 2018;21(2):161-65.
7. Heindel GA, McIntyre CM. Contemporary challenges and novel strategies for health-system formulary management. *Am J Health Syst Pharm*. 2018;75(8):556-60.
8. Joint Commission Resources Portal. M.02.01.01. Edition Hospital Accreditation Program. Retrieved March 29, 2019. Accessed March 7, 2021. <https://e-dition.jcrinc.com/MainContent.aspx>
9. DNV GL Healthcare. National Integrated Accreditation for Healthcare Organizations Interpretive Guidelines and Surveyor Guidance – Version 11. June 17, 2014. Accessed March 7, 2021. <https://www.dnvghealthcare.com/dnv-gl-niaho-standards>
10. Tyler LS, Cole SW, Russell May J, et al. ASHP guidelines on the pharmacy and therapeutics committee and the formulary system. *Am J Health Syst Pharm*. 2008;65(13):1272-83.
11. Malone PM, Malesker MA, Kar I, Glowczewski JE, Fox DJ. Chapter 13: Pharmacy and therapeutics committee. In: Malone PM, Malone MJ, Park SK. eds. *Drug Information: A Guide for Pharmacists*. 6th ed. New York: McGraw-Hill; 2017. Accessed March 7, 2021. <http://accesspharmacy.mhmedical.com/content.aspx?bookid=2275§ionid=177199676>
12. 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(24):2925-34.
13. Cerny 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(8):1547-73.

14. National Comprehensive Cancer Network. NCCN Clinical Practice Guidelines in Oncology with NCCN Evidence Blocks. June 24, 2020. Accessed March 7, 2021. <https://www.nccn.org/evidenceblocks/>
15. Neumann PJ, Willke RJ, Garrison LP, Jr. A health economics approach to US value assessment frameworks-introduction: an ISPOR special task force report [1]. *Value Health*. 2018;21(2):119-23.
16. Sorenson C, Lavezzari G, Daniel G, et al. Advancing value assessment in the United States: a multistakeholder perspective. *Value Health*. 2017;20(2):299-307.
17. Mariotto AB, Yabroff KR, Shao Y, et al. Projections of the cost of cancer care in the United States: 2010-2010. *J Natl Cancer Inst*. 2011;103:117-28.
18. Saunders K, Mably MS, Shull SS, et al. Implementing value assessment in oncology practice: a single-center experience. *J Oncol Pharm Pract*. 2019;25(4):947-53.
19. Studdert AL, Gong GL, Srinivas S, Chin, AL, Deresinski S. Application of pharmacoeconomics to formulary management in a health system setting. *Am J Health Syst Pharm*. 2019;76(6):381-86.
20. Bentley TGK, Cohen JT, Elkin EB, et al. Measuring the value of new drugs: validity and reliability of 4 value assessment frameworks in the oncology setting. *J Manag Care Spec Pharm*. 2017;23(6-a Suppl):S34-S48. Accessed March 7, 2021. <https://www.jmcp.org/doi/abs/10.18553/jmcp.2017.23.6-a.s34>
21. Bentley TGK, Cohen JT, Elkin EB, et al. Validity and reliability of value assessment frameworks for new cancer drugs. *Value Health*. 2017;20(2):200-05.
22. Wilson L, Lin T, Wang L, et al. Evaluation of the ASCO Value Framework for anticancer drugs at an academic medical center. *J Manag Care Spec Pharm*. 2017;23(2):163-69. doi: 10.18553/jmcp.2017.23.2.163
23. Bang H, Zhao H. Median-based incremental cost-effectiveness ratio (ICER). *J Stat Theory Pract*. 2012;6(3):428-42.
24. Anderson JL, Heidenreich PA, Barnett PG, et al. ACC/AHA statement on cost/value methodology in clinical practice guidelines and performance measures. *J Am Coll Cardiol*. 2014;63(21):2304-22.
25. R Foundation for Statistical Computing. The R Project for Statistical Computing. Accessed March 7, 2021. <https://www.r-project.org/>
26. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol Assess*. 1994;6(4):284-90.
27. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155-63.
28. Taylor PJ. An introduction to intraclass correlation that resolves some common confusions. Faculty paper. University of Massachusetts, Boston, MA. 2010. Accessed March 7, 2021. <http://www.faculty.umb.edu/pjt/09b.pdf>
29. Walter E. Approaches to capturing value in oncology. *Recent Results Cancer Res*. 2019;213:85-108.
30. Becker DJ, Lin D, Lee S, et al. Exploration of the ASCO and ESMO Value Frameworks for antineoplastic drugs. *J Oncol Pract*. 2017;13(7):e653-e665.
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(24):2764-71.
32. Slomiany M, Madhavan P, Kuehn M, Richardson S. Value frameworks in oncology: comparative analysis and implications to the pharmaceutical industry. *Am Health Drug Benefits*. 2017;10(5):253-60.
33. McCabe C, Claxton K, Culyer AJ. The NICE cost-effectiveness threshold: what it is and what that means. *Pharmacoeconomics*. 2008;26(9):733-44.
34. Grosse SD. Assessing cost-effectiveness in healthcare: history of the \$50,000 per QALY threshold. *Expert Rev Pharmacoecon Outcomes Res*. 2008;8(2):165-78.
35. Schnipper LE, Bastian A. New frameworks to assess value of cancer care: strengths and limitations. *Oncologist*. 2016;21(6):654-58.
36. Chandra A, Shafrin J, Dhawan R. Utility of cancer value frameworks for patients, payers, and physicians. *JAMA*. 2016;315(19):2069-70.
37. Dubois RW, Westrich K. As value assessment frameworks evolve, are they finally ready for prime time? *Value Health*. 2019;22(9):977-80.