
Richard J. Willke, PhD1,*; Peter J. Neumann, ScD2; Louis P. Garrison Jr, PhD3; Scott D. Ramsey, MD, PhD4

1International Society for Pharmacoeconomics and Outcomes Research, Lawrenceville, NJ, USA; 2Center for the Evaluation of Value and Risk in Health, Tufts Medical Center, Boston, MA, USA; 3The Comparative Health Outcomes, Policy, and Economics Institute, University of Washington, Seattle, WA, USA; 4Department of General Internal Medicine, University of Washington, Seattle, WA, USA

ABSTRACT

The sixth section of our Special Task Force (STF) report reviews and comments on recent US-oriented value assessment frameworks, specifically those published by the American College of Cardiology/American Heart Association, the Institute for Clinical and Economic Research, the American Society of Clinical Oncology, the National Comprehensive Cancer Network, and the Memorial Sloan Kettering Cancer Center. We review published commentaries that address the validity, reliability, and conceptual underpinnings of these frameworks. We find common themes of critique regarding the strengths and limitations across frameworks. Particular shortcomings of some frameworks pose greater threats to their face validity and utility compared with others. The most significant limitations include lack of clear perspective (e.g., patient vs. health plan) and poor transparency in accounting for costs and benefits. We then review how each framework adheres to core STF recommendations, with particular emphasis on whether the framework can be used to support coverage decisions by health insurers, and whether it adheres to core principles of cost-effectiveness analysis. The Institute for Clinical and Economic Research framework most closely adheres to core STF recommendations. Others have significant limitations that vary widely from framework to framework. We also review how the frameworks follow STF recommendations for addressing potentially relevant issues beyond cost-effectiveness analysis - for example, equity in resource allocation and patient heterogeneity. Finally, we review whether and how each framework uses value thresholds and addresses affordability concerns. We conclude with suggestions for further research, particularly in the areas of testing the measurement and use of novel elements of value and deliberative processes.

Keywords: guiding principles, health technology assessment, methodological issues, value frameworks.

Copyright © 2018, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

Introduction

The work of our Special Task Force (STF) complements previous critiques of US value frameworks (e.g., Refs. [1–5]). In this section, we briefly review the relevant frameworks, summarize several previous reviews, and provide suggestions for those frameworks [6–11].

Recent US Value Frameworks

The five recent US value frameworks identified and described in the articles by Neumann et al. [12] and Garrison et al. [13] of our STF report can be broadly characterized in terms of their decision contexts and perspectives, which in turn determine how each framework accounts for costs and health outcomes. The frameworks are diverse in how they define and measure value [1,12,13]. The Institute for Clinical and Economic Review (ICER) and the American College of Cardiology/American Heart Association (ACC/AHA) frameworks use conventional cost-effectiveness analysis (CEA) with quality-adjusted life-years (QALYs) to support population-level recommendations (including payer, health sector, or societal perspective). The ICER framework is oriented toward health plan coverage and reimbursement decisions, and it also applies CEA with QALYs and estimates population-level budget impact to address affordability when applicable. The ACC/AHA framework is oriented toward supporting the development of clinical guidelines and pathways. The American Society of Clinical Oncology (ASCO) and the National Comprehensive Cancer Network (NCCN) frameworks are oriented toward shared decisions by clinicians and patients for cancer drug therapies.

Conflicts of interest: The authors state that no ethical approval was needed. The views expressed in this article are the personal views of the authors and may not be understood or quoted as being made on behalf of or reflecting the position of the agencies or organizations with which the authors are affiliated.

* Address correspondence to: Richard J. Willke, International Society for Pharmacoeconomics and Outcomes Research, 505 Lawrence Square Boulevard S, Lawrenceville, NJ 08648.

E-mail: rwillke@ispor.org

1098-3015/$36.00 – see front matter Copyright © 2018, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

As such, ASCO and NCCN eschew QALYs in favor of disaggregated measures of outcomes, drawing largely on dimensions of benefits and harms widely recognized by the oncology community. The ASCO and NCCN frameworks consider cancer drug costs only, from either the physician or the patient perspective. Memorial Sloan Kettering Cancer Center’s DrugAbacus also focuses on cancer drugs and is geared toward policymakers, and allows users to specify their willingness to pay for health benefits and the importance they place on innovation and development costs. Using this information, the DrugAbacus calculator then determines the user’s ‘value’ price for a drug indication and compares this with the list price.

Among other frameworks, the Avalere/FasterCures Patient-Perspective Value Framework (PPVF) is a “patient-centric” framework that includes a detailed set of elements that may be important to a patient’s treatment decision [14]. In a somewhat similar fashion, the National Health Council’s rubric sets forth a set of principles for evaluation of frameworks from a patient-centric perspective [15]. Other US-based frameworks include those recommended by the First and Second US Panels on Cost-Effectiveness in Health and Medicine as well as others listed in the article by Neumann et al. [12] of this report [11,16,17]. In Europe, many national health technology assessment bodies and the European network for Health Technology Assessment have developed value frameworks [18]. The European Society for Medical Oncology (ESMO) created a value framework on the basis of “magnitude of clinical benefit” that was similar to ASCO’s, also focusing on shared decision making [19].

Given the number and diversity of these frameworks, we will not attempt a comprehensive critique. Moreover, we recognize that the frameworks continue to evolve—even during the drafting of this report, some of the frameworks have updated their approaches. This section focuses on a selected set of general concerns that are most relevant to key value and decision aspects as discussed in foregoing sections.

**Critiques of the Recent US Frameworks**

**Conceptual Underpinnings and Methodological Issues**

Some observers have criticized the new non-QALY-based frameworks for their lack of a conceptual foundation, for omitting key components of cost or outcome given their perspectives, and for not considering uncertainty. For example, Westrich [2] argues that the oncology-oriented frameworks use untested methods, are confusing in their choices for inputs and outputs, are not patient-centered, and/or do not consider system effects beyond the drug therapy. Several authors note that the ASCO framework omits several key value components, does not fully account for costs of care, and uses arbitrary methods to assign value points. The ASCO and ACC/AHA frameworks have also been criticized for not having well-articulated approaches for reflecting uncertainty [1,3,20,21].

**Validity and Reliability**

As with other rating scales, value assessment frameworks should demonstrate validity (the technique should measure what it intends to measure) and reliability (consistency in repeated measures by the same individual of the phenomenon under investigation) [22,23]. In a pilot study, Bentley et al. [4] found that the ASCO, ICER’s evidence rating matrix, ESMO, and NCCN frameworks demonstrated convergent validity and good agreement across a sample of raters in their relative rankings of value. The ASCO and ESMO frameworks showed good inter-rater reliability; that is, different raters gave similar scores for each instrument. The inter-rater reliability for the ICER framework was not as high, perhaps because of the nature of the conceptual framework for its evidence rating matrix, which was not intended to be numerically scored.

Cohen et al. [5] evaluated the face validity of the ASCO, ICER, and DrugAbacus on the basis of four proposed logical criteria. They found several challenges pertaining to the frameworks’ use of value criteria, concluding that although these frameworks “capture value in a way that is important to various audiences, they are not always logical or consistent.” In part, these inconsistencies stem from problems of perspective, for example, whether and how they take the perspective of clinicians or patients.

**Guiding Principles for Value Frameworks**

A number of researchers and organizations have formulated lists of “good principles” that they believe should guide the development and use of value frameworks [15,16,24–28]. The principles summarized in section three, including calls for transparent, explicit approaches for developing frameworks and reports; consideration of patient-centered care and individualized decision making; the inclusion of a broad range of high-quality evidence; the incorporation of the broad effects of interventions such as longer term outcomes; clear statement of the intended use and audience; and the inclusion of a broad range of technologies. Some of the groups have explicitly or implicitly criticized certain frameworks for not adhering to these and other principles. For example, Sorensen et al. [28] observed that the ASCO framework lacked patient engagement in framework design and value assessment, whereas ICER provided only limited access to the economic models underlying their analyses.

**Comparing the STF Recommendations with Recent Value Frameworks**

In offering recommendations regarding value frameworks, we note that population/plan and individual decisions are generally made sequentially—and iteratively over time—and rely on different conceptualizations of value, as discussed in the article by Garrison et al. [13]. In a typical US health insurance setting, plan-level decisions about whether and how to reimburse new technology precede individual clinician-patient decisions about appropriate use of those technologies, with those individual decisions often informed by practice guidelines, clinical pathways, and utilization management practices. (Note that this sequence of decisions may not apply in all cases, e.g., medical devices or ‘medical necessity’ exceptions.) Commercially insured enrollees will have made their enrollment choices before these decisions (if indeed they have a choice of insurance options from their employer), but they may choose to switch plans after a coverage decision has been made. When enrollees become patients, they choose among treatments given physician recommendations and their own preferences for the technology, as well as insurance benefit design considerations (e.g., how much they will pay out of pocket), and their own budget constraints. Importantly, plan-level decisions, such as those pertaining to coverage, tend to rely on population averages, at least within identifiable subgroups, whereas patient-level decisions encompass factors such as individuals’ histories, prognoses, and preferences.

**Core Recommendations**

For plan-level decisions, our STF recommends that CEA be used as a starting point for underlying value constructs from the health insurer perspective. Elements that may be relevant for...
plan-level decisions and that are not typically captured well, if at all, in cost or QALY estimates include productivity (recommended for inclusion in the societal perspective by the Second Panel) and adherence [17]. Other elements that may be relevant include the value of hope, value of insurance, and adjustments for severity of illness to modify utility gains [17]. These other elements may be incorporated via other methods beyond cost-per-QALY analysis, such as net monetary benefit or cost-benefit calculations or MCDA-type decision processes. We encourage frameworks to make allowances for patient heterogeneity as well as for enrollee preferences for health states—to be informed by patient experiences—in plan-level decisions.

Given their use of CEA that is based on cost per QALY, the ICER and ACC/AHA frameworks are consistent with this basic recommendation concerning plan-level decisions. The ASCO, NCCN, and DrugAbacus frameworks are not consistent with it, although importantly, ASCO and NCCN do not claim to be relevant from a plan-level perspective. ICER has also indicated that it will consider productivity and adherence factors in scenario analysis, and consider severity of illness among its “contextual considerations,” whereas ACC/AHA has not explicitly addressed these issues [29].

**Incorporating Other Potentially Relevant Issues**

**Other attributes of treatment**

Health economists and outcomes researchers have not achieved consensus on how best to measure and incorporate more “novel” attributes associated with treatments (whether via health state utilities or as add-ons or modifiers). This situation is reflected in the conventional stance of ICER and ACC/AHA on the use of cost-per-QALY analysis. The ICER framework, however, includes "direct health benefits that are not adequately captured by the QALY” among its contextual considerations. DrugAbacus includes novelty of treatment and population health benefit in its formula. The field has also not agreed on whether a pure cost-per-QALY approach or a multicriteria decision analysis (MCDA) approach is preferable and feasible for aggregating attributes not typically included in cost-per-QALY analysis. Notably, ICER has indicated that it has attempted formal MCDA in the past and “found the technique too complicated for reliable use,” underscoring the importance of optimizing both the specific approach and the tools to enable it if it is to be more broadly used. Its contextual considerations are rated qualitatively and included in discussions of its appraisal committee before voting [29]. Nevertheless, the draft PPVF framework appears to be following an MCDA-like methodology to weight the attributes included in its framework.

**Equity in health care allocation**

The concept of equity is an important but problematic concern for plan-level decisions, particularly because it pertains to persons with rare diseases explicitly, it may do so implicitly to a degree through its general financial risk protection orientation. Incorporating the value of insurance in CEA could help support coverage of certain treatments for rare diseases. In general, however, incorporating equity considerations in value metrics originally developed for efficacy considerations is challenging. MCDA or deliberative processes may be useful alongside traditional value-based approaches in guiding decision makers in this area.

**Uncertainty**

Both the quality of the underlying evidence and its statistical precision have important implications for the creation and interpretation of value. Statistical uncertainty of costs or outcomes because of sampling should be represented appropriately, at a minimum using one-way and multivariable uncertainty analysis in presenting CEA results or other metrics. ICER’s reviews rate the quality of the comparative effectiveness evidence and typically include probabilistic sensitivity analysis in their economic models, as is consistent with good practice standards for CEA. ICER could also recognize the statistical uncertainty in the evidence more explicitly during their subsequent discussions of comparative value and pricing. ACC/AHA recommends use of the cost-effectiveness acceptability curve to represent the uncertainty in the economic evaluation; it also recommends consideration of the quality of the underlying evidence. NCCN includes an evidence block for the quality of the evidence, and ASCO states that only the highest quality evidence be used. Neither ASCO nor DrugAbacus provides formal recognition of statistical uncertainty in their numerical ratings.

**Patient heterogeneity**

No existing value framework accommodates both the plan/population and patient decision-making perspectives to produce a single, broadly applicable decision aid (nor would it be easy to produce one). Insured patients face only a small fraction of the total cost of their care and have limited incentives to consider costs they do not face. Plan-level measures do not fully capture variation in patient factors, such as individual patient characteristics, treatment responses, and preferences except perhaps through analysis of subgroups or explicit modeling of heterogeneity. Because decisions are most often made sequentially—enrollee and patient decisions follow plan-level decisions—plan-level decisions should leave some flexibility to allow for individual and patient heterogeneity, as discussed earlier and in the article by Garrison et al. [13]. ICER recognizes the potential for heterogeneity of treatment effect in its comparative effectiveness analysis, and ACC/AHA discusses the limitations of CEA when data are inadequate to estimate subgroup effects. In part, to recognize and manage such heterogeneity—both observed and unobserved—traditional health plan strategies include previous authorization, step-edits, or an exceptions process, or other channels through which coverage and reimbursement decisions are allowed to vary, typically by request of the physician acting as the patient’s agent. These patient-level restrictions should ideally follow guidelines for good practice (e.g., transparency), such as those outlined in Perfetto et al. [15], Sorensen et al. [28], and Allen et al. [35].

Patient-oriented value constructs can play a useful role within the coverage and reimbursement constraints established by health insurers. For example, within the context of a value-based formulary, where patients may face differential co-pays for different drugs, a patient-oriented value construct could inform patients about the relative benefits and costs of therapies (e.g., branded vs. generic medications for diabetes). The most useful constructs would allow patients to differentiate among options available to them using value elements relevant to their
preferences or the shared decision-making context. Armstrong and Mullins [36], for example, suggest a logical way to group such elements into global, decisional, situational, and external components. Ideally, those elements would need weighting by patient-specific preferences to lead to genuinely individualized decisions. The PPVF covers many of those elements, along with patient-specific weighting; the validation needed is in progress at the time of this writing. Most patient-level decisions do not require a formal framework or decision aid, because there are a limited number of treatment options that vary on only a few dimensions. Such decisions are made routinely in the clinic on the basis of discussions between patients and physicians. Nevertheless, some treatment decisions are more complex and would benefit from a well-constructed and well-validated patient-oriented framework, such as the MCDA approach described in the article by Phelps et al. [32].

**Cost-Effectiveness Thresholds and Affordability Criteria**

**CEA Thresholds**

As discussed in the articles by Phelps et al. [32] and Danzon et al. [37], a value framework, to be useful in decision making, should specify a standard or benchmark measure of value—often called a threshold—that must be met to justify the cost of a treatment. Thresholds could take the form of ranges and to vary across plans or different situations, as discussed hereafter.

The ASCO and NCCN frameworks can be used to compare aspects of different regimens, but do not provide specific decision criteria, presumably with the intent of allowing patient-level decisions to be determined by individual considerations. The DrugAbacus framework, in determining a target price, implies that such a price is the highest acceptable price for the product. In frameworks that use formal CEA, such as in the ICER and ACC/AHA frameworks, specifying a benchmark measure of value means setting an appropriate incremental cost-effectiveness threshold. Both the ICER and ACC/AHA frameworks have cited threshold ranges that they deem acceptable—$50,000 to $150,000/QALY for ACC/AHA and $100,000 to $150,000/QALY (but allowing for cases as low as $50,000 or as high as $175,000) for ICER—depending on the context of the intervention and condition. Those levels are consistent with thresholds most often cited by US researchers, although some have argued for higher cutoffs (e.g., Ref. [38]).

Different concepts have emerged regarding what a cost-effectiveness threshold should reflect and how it can be estimated [17,39]. One approach, sometimes termed a “demand-side” threshold, relates to the willingness of individuals or the public to pay for health improvements (and thus forgo other types of consumption) [17]. In contrast, a “supply-side” threshold would emphasize the opportunity cost associated with allocating a health plan or system’s resources to a particular use. As the Second Panel on CEA has noted, some have argued that demand-side considerations are secondary to the supply-side constraints on resources available to deliver health care [17,39,40]. Some researchers have attempted to measure the health opportunity cost threshold in various countries and have suggested that thresholds at or significantly lower than per-capita gross domestic product may be appropriate [41]. Nevertheless, others have contested this conclusion [42,43].

The appropriate thresholds will presumably change with the decision maker and the constituencies they represent. A challenge for any population-level framework in the United States is heterogeneity among the decision makers who the framework is intending to inform. In the private payer market, there are multiple decision makers each with their own budgets, input prices, and opportunity costs, and each with their own enrollees who have their own preferences. Moreover, for public payers, such as Medicare and Medicaid, the opportunity cost associated with an intervention extends to forgone goods and services outside the health care sector [17]. Furthermore, decision makers in the United States do not explicitly use strict cost-effectiveness thresholds, and CEA is only one of many inputs into decisions. Both ICER and ACC/AHA try to accommodate a flexible threshold within their acceptable ranges with “contextual considerations,” such as disease severity and ethical considerations, but their conceptual basis for making such adjustments is not clear.

**Affordability**

Considering CEA in the context of affordability (patient, payer, or society) is another long-standing issue, with payer and societal considerations, as discussed in the article by Danzon et al. [37]. The ICER framework has included a budget impact analysis to quantify affordability by estimating the short-term (5-year) costs and savings of an intervention from a payer’s perspective. ICER’s recently revised methodology provides a budget “alert” for new treatments that is based on anticipated national gross domestic product growth rates and the number of new drugs approved each year [29]. The calculation of this budget alert ignores the cumulative budget impact of myriad existing items while focusing on the budget impact of a few new technologies. ICER’s updated framework does highlight the importance of identifying low-value services, and states that future reports will include a list of cost-saving measures in the health system in the relevant clinical area [29]. Still, the ICER budget impact approach can be inefficient and discriminatory because it does not consider explicitly the value of the new interventions under investigation, in relation to the value of interventions already being funded. Another concern is that budget impact analyses with short-term horizons and budget alerts calculated on the basis of the number of drugs approved per year may create disincentives for manufacturers to develop high-value and broadly used medications by creating an inverse relationship between the target population size and the cost-effectiveness threshold applied, which is discouraged in the article by Danzon et al. [37]. It may also trigger utilization management that effectively penalizes patients with more prevalent chronic conditions [44]. As argued in the study by Ciarometaro et al. [45], budget thresholds also violate the criterion that the marginal benefit of a product should equal its marginal cost—which should reflect the same underlying opportunity cost on which the threshold is based—for all products. They also do not rely on the use of the private market mechanism whereby insurers could raise premiums to cover new health technology costs in certain plans, and enrollees then decide whether to purchase those plans—although admittedly these decisions would be affected by the tax subsidy to insurance purchase.

An efficient way to address budget constraints would be to reduce spending on, or to replace, technologies with less favorable cost-effectiveness ratios in favor of budget-expanding but more cost-effective technologies. This could be achieved by price reductions on new technologies, by utilization management targeting less cost-effective subgroups of patients, or by disinvestments in less cost-effective treatments. A lower cost-effectiveness threshold could be set (given some uncertainty about pending new treatments and the success of price reduction and other efforts) that would help achieve the needed overall budget [51]. Any new products (including the new budget-expanding technology in question) as well as existing technologies that could be subject to disinvestment could be held to that new standard. Ideally, an affordability strategy should examine the
entire medical care portfolio subjecting all technologies to the same opportunity cost criterion, rather than assuming that budget savings can be achieved by restricting the price or use of technologies that meet the affordability criteria. Barriers to reducing price or to disinvestment include high transaction costs associated with reducing the use of established technologies within health systems and equity concerns if the technologies of interest are the only effective options for patients with specific conditions.

Conclusions

Recent value frameworks have been helpful in drawing attention to criteria important to different stakeholders in health sector decisions. Nevertheless, given different stakeholder preferences and incentives, it is difficult for any single framework to represent all relevant decision contexts—nor would we expect it to. Thus, it is important for any framework to clearly articulate the value construct it represents and the decision context in which it is to be used, and to be well-validated within that construct and context. For plan-level decisions about coverage and reimbursement, frameworks should build upon principles of efficient resource allocation to maximize population health or well-being, with potential allowances for equity considerations and patient heterogeneity. Patient-level frameworks should be tailored to guiding individual treatment decisions so that patients can consider and weight factors most relevant to their preferences and constraints, within the larger constraints imposed by the insurer using plan-level decision rules. In this section, we have provided comments on the five recent US value frameworks that are based on the first five sections of this report as well as other recent critiques and are firmly grounded in best practices in health economics and outcomes research.

Our STF recommendations in the article by Garrison et al. [13] underscore the need for further research to improve how value assessment frameworks represent the net value of health care interventions for decision-making purposes. Although we believe that the use of QALYs within CEA has substantial benefits, we recognize that it also has limitations [46–49]. We do not believe that these criticisms are significant enough to outweigh their demonstrated strengths as a core measure of health benefit, but agree that careful interpretation and further refinement are needed in some situations, as detailed in the article by Lakkakall et al. [31]. Several novel value elements discussed in this report have the potential to address certain limitations and thus be important to estimates of the value of particular technologies; they may be incorporated into the QALY construct itself or into a broader augmented CEA measure, or be included in an MCDA or as part of a deliberative process. Additional studies on how these value elements can be measured in a rigorous and systematic fashion, across both health and non-health-related interventions when relevant, will be important. Equally important will be testing how they can be used in private and public health plan decision making, because the addition of new value elements may affect decision criteria such as cost-effectiveness thresholds, and may also differentially affect various patient populations. Given the need for stakeholder involvement and, in particular, patient engagement in decision making, further work to refine and to improve the methods and ease of use of deliberative processes such as MCDA is also important. At the patient level, construction of parsimonious value frameworks/decision aids, and testing to ensure both validity and feasibility, would be useful. Finally, adaptation and testing of such methods across diverse interventions—drugs, devices, clinical services, public health programs, and so forth—would lead to improved, and indeed potentially transformative, health care decision making for all patients.

Source of financial support

The authors did not receive any funding for this work other than reimbursement from the International Society for Pharmacoeconomics and Outcomes Research for travel expenses, as needed, for two Special Task Force meetings.

References


