ScienceDirect

Contents lists available at sciencedirect.com
Journal homepage: www.elsevier.com/locate/jval

Notions of "Value" in Healthcare

Valuing Health: Evolution, Revolution, Resistance, and Reform

Charles E. Phelps, PhD, MBA[1,*] Guruprasad Madhavan, PhD, MBA[2]

[1]University of Rochester, Rochester, NY, USA; [2]National Academies of Sciences, Engineering, and Medicine, Washington, DC, USA

ABSTRACT

A number of methods have sought to determine the value of interventions and services that promote health, even when no agreement exists on the proper way to determine and define "value." Previous valuation efforts began simply by counting deaths or measuring life expectancy, slowly evolving to the widespread use of cost-effectiveness analysis (CEA) as the de facto normative standard for medical interventions. Users of CEA recognize that the method is incomplete. Further, no meaningful agreement exists on how best to apply CEA in decision settings because of either inadequacies in the CEA framework or lack of consensus on how to use it in a setting with budget constraints. Yet efforts to value health still predominantly use (and continue to recommend) this limited framework. Is this owing to a lack of new ideas and motivation, resistance to change, or an aversion to embrace more comprehensive systems approaches? We argue that tools of systems engineering can advance our capabilities, but they have had only limited use in health policy. We identify some reasons and specifically highlight the promise of systems-analytic platforms—such as multicriteria decision support systems—and the need to make them more accessible for different uses in real situations with real consequences. We also explore the need for comparative testing of different multicriteria approaches (including direct comparisons with CEA) to learn when and by how much the recommendations differ and what the consequences might be.

Keywords: budget and resource analysis, decision support systems, multi-criteria decision analysis, systems engineering

VALUE HEALTH. 2019; 22(5):505–510

Evolution

In his classic 1962 work, historian Thomas Kuhn proposed that scientific progress occurs not steadily, but rather in a succession of ragged steps that may eventually culminate in revolutionary shifts to our understanding and analysis.[1] Most scholars engage in "normal" science and conduct incremental (or even lesser) research by following received principles and protocols. A few change agents uproot standard methods by finding crucial tests that undermine their applicability. A new "paradigm" then takes hold, and normal science embarks on a slow transition from the old to the new until another puzzle emerges that cannot be solved by the existing mindset but requires another "scientific revolution." We believe that the pathway to determining the "value" of health has followed a similarly jagged course that will likely continue into the future.

A brief history of the ways in which we have determined value of health and medicine demonstrates this phenomenon. The earliest and easily measured concept of prevented premature deaths (erroneously defined as "lives saved") rapidly evolved to the concept of "life-years saved" (by blending "lives saved" with actuarial tables for expected survival by age). The related economic value was initially measured by their foregone workforce productivity.[2] The incompleteness of this approach soon became apparent, and insights from the emerging paradigm of "household production economics"[3,4] shortly led to value measures based on well-being (utility) arising from better health.[5]

The concept of adjusting life years by their quality followed as a development,[6,7] later formalized using a multivariate framework.[8] The concept of quality-adjusted life years (QALYs) was developed as a unified population-based measure of health.[9,10] A subsequent, parallel concept of disability-adjusted life years (DALYs) also achieved widespread application, especially in developing countries. While originally relying on expert-estimated adjustments to life years, current applications of DALYs use population-based assessments akin to those used in QALYs.[11] Another method to describe the health-related quality of life uses five descriptive dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety or depression. Patient rating of these attributes on five levels results in a combination of numbers that summarize the respondent's health state.[12]

Mathematical proof that CEA (using QALYs) derives from maximizing expected utility followed well after QALYs was developed over 20 years ago, supporting its continued and growing use.[13] The incremental cost-effectiveness ratio (ICER)—defined as ratio of incremental costs to incremental QALYs gained—has become the *de facto* normative standard for evaluating the performance and desirability of medical interventions wherever formal models are used for health technology assessment. For example, encouraged by the guidance of economist Alan

* Address correspondence to: Charles E. Phelps, PhD, MBA, University of Rochester, Gualala, CA 95445, USA. Email: charles.phelps@rochester.edu

https://doi.org/10.1016/j.jval.2019.01.010

Williams, CEA has been used to supply policy and investment advice in the British National Health Service at least since 1999, with the founding of their National Institute for Health and Care Excellence (NICE).[14] Its use and influence have steadily expanding since then.[15]

Making decisions based on CEA requires the choice of a maximum willingness to pay (WTP) per QALY. A vigorous debate that has spanned decades still continues on this matter. Initial efforts using "league tables" assessed the cost per QALY for well-established medical interventions to provide guidance on maximum WTP.[16] This approach inevitably leads to arbitrary choices because the appropriate boundary between "reasonable" and "unreasonable" is not definable in natural terms. Further, the league table approach ignores the effect of health insurance on expanding the use of medical care,[17] creating a natural upward bias in the estimated maximum WTP.

A different approach seeks to determine the value of a statistical life by assessing wage premiums for risky occupations such as fire response and police. Nevertheless, that approach contains a major upward bias because it assesses the willingness to accept risk (WTA) rather than the WTP to avoid it. A direct comparison of these two approaches—when possible—shows that the willingness to accept risk approach creates an estimated fivefold (or more) upward bias in the estimates of the conceptually desired WTP measure.[18] Other difficulties with the value of a statistical life (VSL) literature in addition to the WTA/WTP problem are discussed elsewhere.[19,20]

Separately, several organizations have promulgated advice on the maximum WTP. These principally include the World Health Organization (WHO), which has recommended that its member states use 1X to 3X per-capita GDP as the relevant cutoff—1X being a very good investment and 3X being the maximum WTP.[21,22] Some organizations and analysts in the United States suggest a range of $50 000 to $150 000 per QALY.[16,23] Recent WHO publications have cautioned against the use of specific cost-effectiveness thresholds, but do not offer guidance on alternatives to their use.[24,25]

Most recently, a way to estimate the optimal WTP cutoff using utility function parameters such as relative risk aversion has been proposed that does not rely on league tables or wage differentials.[20] These estimates place the optimal WTP cutoff at approximately twice the income for a representative individual with income equal to average GDP for an advanced economy such as the United States, United Kingdom, or many European Union nations. Applying this approach to determine the best single cutoff for a universal healthcare plan requires assessing a number of additional factors. They could include whether or not private supplementation is allowed and whether the choice is made by a welfare-maximizing model or by popular vote using median voter's preferences.

As CEA became the *de facto* normative standard to evaluate medical interventions, the conduct of "normal" science in this space has resulted in thousands of studies, with one repository[26] containing more than 6500 (almost 7300 and counting …) published estimates of cost-effectiveness ratios. Several specialized technical journals exist that focus almost entirely on the methodologies and estimates of CEA. The recent announcement in the United States by CVS Caremark to use a formal $100 000 per QALY cutoff in its drug formulary was the first of its kind—made all the more important with federal approval of the merger between CVS and Aetna, a large health insurer.[27] In the United Kingdom, the NICE recommendations have increased the CEA approval cutoff (making it more generous) for rare diseases, cancer treatments, and end-of-life care. Despite these developments, it is widely understood that CEA is incomplete. It cannot incorporate issues involving comparisons across different members of communities—fundamentally, differential weights that consider inequities or disparate outcomes for different population subgroups. Doing that requires something akin to a "societal utility function" that includes the well-being of individuals, each possibly with different utility weights.[28]

CEA has several other shortcomings. Many ways are available to estimate the required quality adjustments—the conceptually preferred but difficult to administer "standard gamble" or easier-to-administer time-trade-off and "visual analog" approaches.[29] Other methods include "conjoint analysis," a tool with origins in marketing research[30] and discrete choice experiments.[31] Each of these can yield different results and create different difficulties in implementation. Nevertheless, none of these approaches can meaningfully capture all of the potential health states—and their associated complexities—that decisively affect real decisions in health and medicine.[32,33] As an illustration, the use of CEA in pharmacy and therapeutics (P&T) committees in hospitals and health plans have produced mixed results, thus limiting its use.[34–36]

Conducting "quality adjustment" sometimes requires a population-based measure of an individual's quality of life, and this may include individuals with a plethora of complex chronic debilitations. Typically, these adjustments are carried out using multiplication or addition of numerous single condition measures, which indicates that realistically we are not likely to have estimates of quality adjustment for many situations that importantly affect people's lives. Thus, this challenge remains outside the realistic capabilities of CEA.

When using CEA, people often deal with the complexity by ignoring other attributes beyond QALYs. But this does not really solve anything—it merely masks the existence of a problem. One recent proposal focuses on "expanded" CEA (ECEA) that employs available data on financial risk as well as equity and distribution in a decision tree structure.[37] Another effort has identified 10 value attributes beyond QALYs, known as "augmented" CEA.[32] Besides financial risk and equity, the augmentation extends to variables such as fear of contagion, risk, hope, option value, and productivity enhancement. Nevertheless, neither ECEA nor augmented CEA provide a sound basis for investment or policy decisions because they do not create a composite figure of merit useful to compare alternative interventions. Augmented CEA has the potential for providing a composite measure, but that would require measuring every possible dimension of value either in monetary units or QALYs, a task that remains beyond current methodological approaches. Even if successful, this approach could raise similar objections to those originally used in rejecting cost–benefit analysis (CBA) and shifting to CEA, which measures health benefits in natural units.

This leads to our first question: Despite widespread understanding that CEA methods are incomplete, why do people continue using them and recommending their further use? Simply put, analysts and decision makers seem comfortable (even complacent) with CEA, despite its known limitations, and even resistant to other viable new ideas, perhaps because of its perceived value across different situations. Therefore, no alternative approach has become widely accepted, thus anchoring CEA as a mainstream practice.

Revolution

It is a basic point that interventions and services in health and medicine have multiple dimensions of value. Therefore, the concept of value, in a very subjective sense, relies on what

perspective an individual or an entity brings to it, and the context from which value is thought to emerge. An approach to advance our thinking about value has been available for decades, rooted in the practical traditions of systems engineering, operations research, and decision theory—multicriteria decision analysis (MCDA).

In recent decades, MCDA has been considered a viable subject of exploration in the realm of health and medical policy. Earliest efforts arose from decision theorists, and the first decision support software (ELECTRE) appeared in the 1960s.[38] With a solid theoretical and operational foundation,[39] the practice of MCDA has achieved wider use in other fields such as nuclear waste planning,[40] environmental policy, and urban planning. In the United States, an MCDA application appeared in the realm of healthcare technology evaluation in 1988,[41] slowly followed by other uses,[42] and real influence on decision making has yet to be reported. In 2016 International Society for Pharmaceutical Outcomes and Research (ISPOR) task force reports[43,44] reviewed the rapidly growing use of MCDA for health policy applications, which included early examples of planning for Canadian health regions in 2003[45] and in the biopharmaceutical sector, beginning in 1998.[46]

MCDA comes in a variety of flavors, including the early entrant (in English translation) ELimination and Choice Expressing Reality (ELECTRE), Multi-Attribute Utility Theory (MAUT), Analytic Hierarchy Process (AHP), Measuring Attractiveness by a Categorical Evaluation Technique (MACBETH), Preference Ranking Organization METHod for Enrichment of Evaluations (PROMETHEE), and a panoply of other acronym-rich approaches. The International Society for Multi-Criteria Decision-Making lists over 2 dozen available software implementations,[47] with each MCDA approach differing in an important way from standard economic models. Instead of *estimating* the relevant parameters from observed behavior (economic models typically measure price and income effects on demand), the models *elicit* the preferences of real people involved in decision making.

MCDA models typically create a composite measure that is a weighted sum of partial-value functions. Each model differs in ways user preferences are elicited and converted into utility weights, and in setting attribute scales for analyses. Thus, MCDA approaches can bypass the limitations of CEA and offer the following merits,[48] including:

- Improving transparency
- Guiding and refining data collection and information gathering
- Reverse engineering of product design in multidimensional preference space
- Bypassing human cognitive errors, especially probabilistic factors
- Supporting decision convergence
- Allowing scenario test drives before making real investments

To elaborate briefly on the question of transparency, a principal benefit of MCDA is that it makes tradeoffs across attributes explicit. With CEA, researchers regularly state (or should admit) that their analysis is incomplete, and other factors affect real-world decisions. However, they leave these for decision makers to incorporate in practice, and that step is where transparency typically evaporates, with final decisions resting on some unstated combination of CEA ratios and the "other issues."

Expanding insights from social and behavioral sciences have shown major limits to human decision making, especially with overestimating and overweighting low-probability events.[49] Framing of issues also looms large. As an illustration, in an experiment at Harvard University, samples of patients, graduate students, and physicians all preferred different treatment options depending on whether data about surgery outcomes were presented as survival rates (more favorable towards surgery) or mortality rates (less favorable towards surgery), even though the data contained identical information about risks.[50]

This leads us to our second question: Despite these advantages and demonstrated track record in other areas of consequence, why has MCDA not come into more widespread use in health and medicine? The reluctance to test and adopt MCDA may be linked to potential misconceptions about it, the cost of implementation, lack of replicability of results as participants change, or its sheer comprehensiveness (the very advantage being perceived as a threat), or simply the lack of willingness to use or experience using it.

Resistance

Based on our experiences in developing an MCDA-based systems analysis platform for prioritizing vaccine development[51] and widely testing and disseminating that tool in numerous international settings, we believe that the current concerns about the use of MCDA typically fall into six categories. We discuss them in turn, along with ways to address them.

"MCDA requires too much data." MCDA indeed requires more data than CEA. For each extra value dimension beyond QALYs (the staple of CEA), data or estimates are required in MCDA. Nevertheless, MCDA does not create this complexity—the nature of problem does. MCDA helps organize the data into valuable information. One cannot meaningfully think about these problems without knowing how each option performs across different attributes. CEA masks this complexity through its narrow approach. Although augmented CEA may point toward or gather the relevant data, it does not yet assemble them into a unified measure until every attribute is measured in QALYs or monetized. Until that point is reached, augmented CEA is not a viable alternative to MCDA.

As recommended by the ISPOR Task Force on US Value Assessment Frameworks,[32] health benefits (eg, QALYs) will continue to be an integral factor, but seldom—if ever—be the only attribute in the valuation of health and medical or population health interventions.

"**MCDA models demand too much from decision makers. They are unwieldy, perhaps intractably so, for collective decisions."** MCDA requires that decision makers specify their preferences and systematically conduct trade-off analyses. These steps require time and thought that any serious decision scenario would entail. The trade-offs involve attributes that are ultimately subjective, such as likelihood of a scientific breakthrough of value in other medications or the acceptance of an intervention within local culture. Quantifiable attributes such as QALYs, disease burden, productivity, or cost data might come from external data or estimation models.

Further, most MCDA models implicitly presume one decision maker, and the decision analysis process may become complicated when the decision maker is actually a group—perhaps even the society at large, as it should be on democratic matters of public policy.

To elicit the values, and then to score candidates being analyzed, a method like AHP requires response to two types of questions.[52] First comes a series of questions for each paired attribute (which is more important, and then, on a 1 to 9 scale, by how much). For N attributes, there are $N(N-1)/2$ pairwise comparisons, and hence $N(N-1)$ decisions. Next, decision makers must evaluate each pair of candidates on each subjectively measured dimension of value, again using two questions (which is

better and relatively by how much). For K candidates and N attributes, this leads to another N(N − 1) question for K(K − 1)/2 pairs of candidates.

Comparatively, with MAUT, to establish the weights decision makers need to thoughtfully provide a rank order of N attributes, thus requiring N − 1 decisions.[53,54] In the most simplified form, analysts or decision makers would then estimate the performance of each candidate across each attribute, requiring N×K decisions. Thus, the decision demands are smaller by a factor of N for weighting and (N − 1)(K − 1)/2 for candidate performance measures in MAUT compared with AHP. Other MCDA approaches generally fall in between MAUT and AHP in terms of complexity.

"You can easily game the system with MCDA and produce any outcome you want." In truth, almost any outcome can be reached by (strategically) altering MCDA weights. Nevertheless, this concern is misplaced. First, in many settings, decision makers have good incentives to reveal their "sincere" preferences. More importantly, the realistic alternative is that many other attributes are brought into play to justify decisions in unknown ways. It is thus easier to get any outcome one might wish when discussion of how other factors entered the decision remains hidden. Formally identifying the attribute weights increases transparency and thus makes it more difficult to arbitrarily reach a predetermined choice. The transparency ensured by MCDA actually reduces the ability of users or decision makers to manipulate the outcome compared with real world alternatives.

"The meaning of the composite MCDA index is obscure. What does it mean?" In MCDA, each composite value index (score) differs from others because the weights differ from user to user. In addition, some MCDA indexes are interval scales that have no natural zero, just as with the temperature scales of Fahrenheit and Celsius. In these interval scales, a value of (say) "3" is not three times as good as a value of "1," it is better by "2" units of value, but those units have no natural meaning. Therefore, one cannot meaningfully take ratios of interval scale cost per unit of value, a standard practice in CEA.

Two remedies deal with this. The first follows the standard CEA approach by using differences in health benefits and costs between the standard intervention and the option under consideration. ICER is based on the difference in costs over difference in health benefits. Because these differences have a natural zero, the ratio has consistent meaning. The second remedy comes from using specific versions of MCDA that create ratio scales. The AHP index is naturally a ratio index, and so 2 is twice as good as 1. Further, available modifications to the MAUT also have a ratio scale interpretation, thus allowing an intuitive interpretation of the meaning of the value index that users can naturally understand.[55]

"MCDA is useless under budget constraints and can't guide resource allocation." This was once a valid criticism, but now at least three approaches have been developed to guide resource allocation using MCDA. To do this, all candidates need to be evaluated using the same multi-attribute utility index (MAUI).

- If an investment budget is set exogenously, sequentially invest in projects, beginning with the most favorable MAUI-to-cost ratio, and continue until the available investment budget is expended.[46] The overall value of portfolios can also be evaluated in this approach by comparing "packages" of investments. Of course, this approach gives no guidance about how to formulate the budget initially.
- If an investment budget is to be determined, a starting point might be to have the relevant population vote on or disclose their willingness to pay for each MAUI unit or an overall investment budget. Each voter can state their WTP per MAUI unit, and a median value can be used as the decision cutoff.[56,57] This approach has been used, for example, in setting property tax rates to finance public schools.[58]
- Another approach to determine the investment thresholds is to extrapolate from previously determined WTP values for QALYs.[19] In this approach, the MAUI would contain QALYs as a major component. Collapsing all other attributes into a single measure (X):

$$\mathrm{MAUI} = w \times \mathrm{QALY} + (1-w) \times X.$$

Then if the WTP for QALYs has been established as K, then the appropriate WTP for MAUI value equals K/w. Thus, for example, if K = $100 000 and w = 2/3, then the WTP for MAUI units equals $150 000.

In all of these approaches, best practices include only true measures of value in the MAUI model (not costs), and then calculate cost per MAUI ratios in the same way that cost per QALY measures separate value and costs in CEA.[44,59]

A concern about CEA (pertaining also to MCDA) is that decisions using approved cutoffs (eg, $100 000 per QALY) may lead to approval of interventions that are otherwise unaffordable. This issue—what is cost-effective may not necessarily be affordable—emerged prominently regarding treatment for hepatitis C. Fully expanding the hepatitis C treatment—or introducing new technologies such as "multi-omics"—could overwhelm state Medicaid systems in the United States.[60] Because the cutoff value and the budget cannot be set independently, one implies the other. A preferred solution might be to bring the two into alignment, either by reducing the announced WTP standard or increasing the budget, or some combination thereof.[59] This issue emerges more forcefully in MCDA, where some of the benefits (value) may not be the responsibility of the relevant health or medical agency, but may fall more into the realms of social welfare, planning for research, development, higher education and even national security. Who should pay for these extra benefits and how that might be accomplished in standard budgetary process (filled with negotiations) remains unclear.

"Wouldn't it be simpler just to 'monetize' all the attributes and include them in the QALY measure?" This is not as straightforward as it may sound and should not necessarily be the approach of choice. Nevertheless, we anticipate that efforts will be undertaken to convert some of these attributes either into dollars or QALYs. A recent ISPOR panel also recognized these possibilities—and associated tools for them—but also concluded that some key attributes will not be able to achieve this goal, including numerous equity and ethical issues, fear of contagion, and likelihood of scientific breakthrough.[32] An approach to monetize the value of insurance and the value of hope has been proposed, but not yet tested.[61]

Reform (and Responsibility)

Medical care costs continue to increase in absolute size and share of the overall economy, rapidly approaching 20% of the total GDP spending in the United States. Health outcomes in the United States are far below those in other nations, despite our spending more for medical care both in dollar terms and as a proportion of GDP than any other nation in the world.[62] Major inequities regarding access to healthcare and health outcomes persist that our current approaches such as CEA cannot resolve. The problem expands even more when the wider domain of population health comes into play.[63] As an oft-repeated saying suggests, insanity is doing the same thing over and over again but expecting different results. This applies to the ongoing use of CEA to rationalize our

healthcare systems. CEA is expected to remain an important staple of health economic discussions, but its intrinsic limitations and inadequacies in capturing other important factors are gaining increasing recognition.

In this regard, no MCDA—or any systems engineering method at large—perfectly resolves all the universal concerns of health and medicine, but we believe that striving for absolute perfection instead of capitalizing on the available good and comprehensive systems approaches could be a liability on multiple fronts. To restate Sulvio Funtowitz and Jerome Ravetz,[64] we live in a "post-normal" world where facts are uncertain, values in dispute, stakes high, and decisions urgent. These are the sorts of the situations where MCDA and related systems-analytic models can operate most beneficially. Thus, a multipronged strategy to improve the use of MCDA and related tools would include the following.

1. Expanding the testing and resulting improvements of MCDA and other comprehensive tools of systems engineering to understand how they operate in consequential scenarios of health and medicine. This would include direct comparison of MCDA with CEA to learn when and how they differ, and what the implications are across different settings.
2. Creating commonly accessible data resources[65] and expanding the number of people with relevant skills to conduct such assessments and comparisons, a task with possible roles for students in schools of public policy, public health, business, engineering, design, and communications.[66] Initiating large research and implementation projects to further refine our understanding and use of MCDA, including complexity of implementation, verification and validation of results, use in group settings, and use in a suite of tools for large-scale monitoring and decision analysis.[63] Additional issues to be considered should include analyses of the "human factors" associated with MCDA: ease of use; ability of users to comprehend how the model works and trust its output; and reliability of results.
3. Understanding how different voting methods perform when used to implement MCDA models in group settings. This would require direct comparison of voting models in various settings.[67,68] Different steps in implementing MCDA models may require different voting methods. These steps would help improve MCDA models, make them more user-friendly, and hence relevantly expand their use for the benefit of population health.

In conclusion, we believe that comprehensive systems-analytic tools such as MCDA models have much benefit to offer in both technical abilities and human factors that can be readily refined further through field experiments and testing. For those who believe that MCDA is not ready for use or expansion in real-world policy analysis, we note that the first consumer automobiles in human society were Model T Fords, not Ferrari race cars. The first Wright Brothers flight in its entirety (120 feet) barely surpassed the wing span of a Boeing 737 airliner, but their fourth flight that day spanned 852 feet, a 7-fold increase from the first flight. In engineering, such developments steadily happen with persistence, improvements, and accountability—precisely the attributes beneficial for making comprehensive and serious decision making with potential to affect the health of populations at large.

As discussed earlier, some key attributes of MCDA-informed value fall into the realm of externalities—as in costs borne or benefits received by society in aggregate that are not captured by individual decision making.[69] As an example of a consideration, the structure and incentives of patent law collide with rapidly expanding prescription drug insurance, jointly affecting the affordability and availability of medicines.[70] Only a comprehensive MCDA model is currently able to integrate these kinds of issues by overcoming the limitations of narrower, reductionist techniques such as CEA. Perfection is not necessary to begin expanding the use of MCDA across decision scenarios, but exposing analysts and decision makers to the merits of comprehensive systems approaches seems vitally essential now. Or, as Buckminster Fuller captured it: "You never change things by fighting against the existing reality. To change something, build a new model that makes the old model obsolete."

Disclosure

The views expressed in this article are those of the authors and not necessarily of the National Academies of Sciences, Engineering, and Medicine.

REFERENCES

1. Kuhn TS. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press; 1962.
2. Mushkin S. Investment in human beings. *J Polit Econ*. 1962;70(5):129–157.
3. Becker GS. A theory of the allocation of time. *Econ J*. 1965;75(299):493–517.
4. Lancaster KJ. A new approach to consumer theory. *J Polit Econ*. 1966;74(2):132–157.
5. Grossman M. On the concept of health capital and the demand for health. *J Polit Econ*. 1972;80(2):223–255.
6. Klarman HE, Francis JO, Rosenthal GD. Cost-effectiveness analysis applied to the treatment of chronic renal disease. *Med Care*. 1968;6(1):48–54.
7. Torrance GW. Measurement of health state utilities for economic appraisal: a review. *J Health Econ*. 1986;5(1):1–30.
8. Pliskin JS, Shepard DS, Weinstein MC. Utility functions for life years and health status. *Oper Res*. 1980;28(1):206–224.
9. Klarman HE. The road to cost-effectiveness analysis. *Milbank Q*. 1982;60(4):585–603.
10. Weinstein MC, Torrance G, McGuire A. QALYs: the basics. *Value Health*. 2009:S5–S9.
11. Saloman JA, Vos T, Hogan DR, et al. Common values in assessing health outcomes form disease and injury; disability weights measurement study for the Global Burden of Disease Study 2010. *Lancet*. 2012;380:2129–2143.
12. Oppe M, Devlin NJ, van HB, Krabbe PF, de CF. A program of methodological research to arrive at the new international EQ-5D-5L valuation protocol. *Value Health*. 2014;17(4):445–453.
13. Garber AM, Phelps CE. Economic foundations of cost-effectiveness analysis. *J Health Econ*. 1997;16(1):1–31.
14. National Center for Health and Care Excellence. History of NICE. https://www.nice.org.uk/about/who-we-are/history-of-nice. Accessed March 10, 2019.
15. Mason A, Towse A, eds. *The Ideas and Influence of Alan Williams*. Oxford: Radcliffe Publishing; 2008.
16. Neumann PJ, Cohen JT, Weinstein MC. Updating cost-effectiveness—the curious resilience of the $50,000-per-QALY threshold. *N Engl J Med*. 2014;371:796–797.
17. Newhouse JP. *Free For All? Lessons from the RAND Health Insurance Experiment*. Cambridge, MA: Harvard University Press; 1996.
18. Tuncel T, Hammitt JK. A new meta-analysis on the WTP/WTA disparity. *J Environ Econ Manage*. 2014;68(1):175–187.
19. Phelps CE, Madhavan G. Resource allocation in decision support frameworks. *Cost Effective Res Alloc*. 2018;16:48.
20. Phelps CE. *Optimal Willingness to Pay in Cost-Effectiveness Analysis, Working Paper*. Rochester: University of Rochester; May 26, 2018.
21. Tan-Torres Edjerer R, Baltussen T, Hutubessy A, et al. *Making Choices in Health: WHO Guide to Cost-Effectiveness Analysis*. Geneva: World Health Organization; 2003.
22. World Health Organization. *Choosing Interventions That Are Cost-Effective*. Geneva: World Health Organization; 2014. http://www.who.int/choice/en/. Accessed March 10, 2019.
23. Institute for Clinical and Economic Review. Final Value Assessment Framework for 2017-2019. https://icer-review.org/final-vaf-2017-2019/. Accessed March 10, 2019.
24. Marseille E, Larson B, Kazi DS, et al. Thresholds for the cost-effectiveness of alternative interventions: alternative approaches. *Bull World Health Organ*. 2015;93:118–124.
25. Bertram MY, Lauer A, De Joncheere K, et al. Cost-effectiveness thresholds: pros and cons. *Bull World Health Organ*. 2016;94:925–930.

26. CEA Registry. Center for the Evaluation of Value and Risk in Health, Tufts Medical Center. https://cevr.tuftsmedicalcenter.org/databases/cea-registry. Accessed March 10, 2019.
27. Abelson R. CVS Health and Aetna $69 billion merger is approved with conditions. New York Times. https://www.nytimes.com/2018/10/10/health/cvs-aetna-merger.html. Accessed March 10, 2019.
28. Bator FM. The simple analytics of welfare maximization. *Am Econ Rev*. 1957;47(1):22–59.
29. Neumann PJ, Sanders GD, Russell LB, et al. *Cost-Effectiveness in Health and Medicine*. 2nd ed. Oxford: Oxford University Press; 2017.
30. Green P, Srinivasan V. Conjoint analysis in consumer research: issues and outlook. *J Consum Res*. 1978:103–123.
31. McFadden DL. Econometric analysis of qualitative response models. In: Griliches Z, Intriligator MD, eds. *Handbook of Econometrics, Volume II*. Chapter 24, Amsterdam: Elsevier; 1984.
32. Lackdawalla DN, Doshi JA, Garrison LP, et al. Defining elements of value in health care—a health economics approach: an ISPOR Special Task Force Report [3]. *Value Health*. 2018;21:131–139.
33. Phelps CE, Lakdawalla DN, Basu A, et al. Approaches to aggregation and decision making—a health economics approach: an ISPOR Special Task Force Report [5]. *Value Health*. 2018;21:146–154.
34. Aspinall SL, Good CB, Glassman PA, Valentino MA. The evolving use of cost-effectiveness analysis in formulary management within the Department of Veterans Affairs. *Med Care*. 2005;43(7 Suppl):20–26.
35. Shulkin D. Reinventing the pharmacy and therapeutics committee. *Pharm Therapeut*. 2012;37(11):623–649.
36. Wang Z, Salmon JW, Walton SM. Cost-effectiveness analysis and the formulary decision-making process. *J Managed Care Pharm*. 2004;10(1):48–59.
37. Verguet S, Kim JJ, Mamison DT. Extended cost-effectiveness analysis for health policy assessmenet: a tutorial. *Pharmacoeconomics*. 2016;34:913–923.
38. Roy B. Classement et choix en présence de points de vue multiples (la méthode ELECTRE). *La Revue d'Informatique et de Recherche Opérationelle (RIRO)*. 1968;8:57–75.
39. Keeney R, Raiffa H. *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. New York: Wiley; 1976.
40. Keeney R. An analysis of the portfolio of sites to characterize for selecting a nuclear repository. *Risk Anal*. 1987;7(2):195–218.
41. Dolan JG. The potential impact of a decision support system based on the analytical hierarchy process (AHP) on physicians decisions regarding antibiotic-treatment. *Clin Res*. 1988;36(3):A710.
42. Köksalan M, Wallenius J, Zionts S. *Multiple Criteria Decision Making: From Early History to the 21st Century*. Singapore: World Scientific; 2011.
43. Thokala P, Devlin N, Marsh K, et al. Multiple criteria decision analysis for health care decision making—an introduction: report of the ISPOR MCDA Emerging Good Practices Task Force. *Value Health*. 2016;19(1):1–13.
44. Marsh K, IJzerman M, Baltussen R, et al. Multiple criteria decision analysis for health care decision making—emerging good practices: report 2 of the ISPOR MCDA Emerging Good Practices Task Force. *Value Health*. 2016;19:125–137.
45. Mitton CR, Donaldson C, Waldner H, Eagle C. The evolution of PBMA: towards a macro-level priority setting framework for health regions. *Health Care Manage Sci*. 2003;6:263–269.
46. Phillips LD, Bana e Costa CA. Transparent prioritisation, budgeting and resource allocation with multi-criteria decision analysis and decision conferencing. *Ann Oper Res*. 2007;154:51–68.
47. International Society on MCDM. Multiple Criteria Decision Making. https://www.mcdmsociety.org/content/software-related-mcdm. Accessed March 10, 2019.
48. Phelps CE, Madhavan G. Using multicriteria approaches to assess the value of health care. *Value Health*. 2017;20:251–255.
49. Kahnemann D. *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux; 2011.
50. McNeil B, Pauker SG, Sox HC, Tversky A. On the elicitation of preferences for alternative therapies. *N Engl J Med*. 1982;306:1259–1262.
51. Madhavan G, Phelps CE, Rappuoli R, et al., eds. *Ranking Vaccines: Applications of a Prioritization Software Tool*. Washington, DC: The National Academies Press; 2015.
52. Saaty TL. *Decision Making for Leaders: The Analytic Hierarchy Process for Decisions in a Complex World*. Pittsburgh, PA: RWS Publications; 2008.
53. Edwards W, Barron FH. SMARTS and SMARTER—improved simple methods for multi-attribute utility measurement. *Organ Behav Hum Decis Process*. 1994;60:306–325.
54. Barron FH, Barrett BE. Decision quality using ranked attribute weights. *Manage Sci*. 1996;42:1515–1523.
55. Phelps CE, Madhavan G. *Scale Me If You Can: Making Multi-Criteria Decision Support Systems User-Friendly*. University of Rochester Working Paper; 2018.
56. Galton F. One vote, one value. *Nature*. 1907:414.
57. Galton F. Vox populi. *Nature*. 1907:450–451.
58. Balinski M, Laraki R. *Majority Judgment: Measuring, Ranking and Electing*. Cambridge, MA: MIT Press; 2010.
59. Garrison LP, Neumann PJ, Willke RJ, et al. A health economics approach to US value assessment frameworks—summary and recommendations of the ISPOR Special Task Force Report [7]. *Value Health*. 2018;21:161–165.
60. Augustine N, Madhavan G, Nass S, eds. *Making Medicines Affordable: A National Imperative*. Washington, DC: The National Academies Press; 2018.
61. Lakdawalla DN, Phelps CE. *Evaluationi of Medical Technologies with Uncertain Benefits*. University of Southern California Price School of Public Policy, Working Paper, March 25, 2019.
62. Phelps CE, Parente ST. *The Economics of US Health Care Policy*. New York: Routledge Press; 2017.
63. Madhavan G, Phelps CE, Rouse WB, Rappuoli R. Vision for a systems architecture to integrate and transform population health. *PNAS*. 2018;115:12595–12602.
64. Funtowicz SO, Ravetz JR. Uncertainty, complexity and post-normal science. *Environ Toxicol Chem*. 1994;12:1881–1885.
65. Madhavan G, Phelps CE, Sangha K, et al. Bridging the gap: need for a data repository to support vaccine prioritization efforts. *Vaccine*. 2015;33S:B34–B39.
66. Madhavan G, Phelps CE, Rappuoli R, et al. Strategic planning in population health and public health practice: a call to action for higher education. *Milbank Q*. 2016;94:109–125.
67. Madhavan G, Phelps C, Rappuoli R. Compare voting systems to improve them. *Nature*. 2017;541:151–153.
68. Madhavan G, Phelps CE. Human factors in democracy. *Bridge*. 2018;48(4).
69. Coase R. The problem of social cost. *J Law Econ*. 1960;3:1–44.
70. Phelps CE, Madhavan G. Patents and drug insurance: clash of the titans? *Sci Transl Med*. 2018;10. pii: eaat6902.