

Moving the QALY Forward: Rationale for Change

Marilyn Dix Smith, BSPharm, PhD,¹ Michael Drummond, DPhil,² Diana Brixner, BSPharm, PhD³

¹International Society for Pharmacoeconomics and Outcomes Research, Lawrenceville, NJ, USA; ²University of York, York, UK; ³University of Utah, Salt Lake City, UT, USA

Keywords: consensus development, EQ-5D, Health Utilities Index, HUI, NICE, preference-based measures, QALY, quality-adjusted life-years utilities.

In November 2007, ISPOR convened the “Building a Pragmatic Road: Moving the QALY Forward” Consensus Development Workshop where thought leaders presented and discussed the following: 1) defining and refining the quality-adjusted life-year (QALY) by analyzing the underlying assumptions; 2) developing alternative paradigms and measures; 3) retaining the QALY and enhancing it with additional methods; and 4) incorporating a variety of decision-makers’ needs. The goal of this workshop on the QALY was to reach consensus among the thought leaders on these broad categories, with recommendations for specific, applicable actions. Workshop participants are listed below.

Before this workshop, participants were divided into five groups and each group met via teleconference to develop their arguments on one of the following topics: 1) defining the QALY; 2) considering alternate definitions of the QALY and defining the issues; 3) incorporating an experience-based approach to the QALY; 4) defining the importance of bridging past QALY information to future QALY information; and 5) defining the value of the QALY from a health-care decision-makers’ perspective.

Using a modified “forming–storming–norming–performing” model [1] for consensus development during the 2-day workshop, each group presented arguments about the QALY, received feedback on their arguments and modified, if needed. Then a group of representatives from each group was convened to develop consensus statements on “moving the QALY forward” based on the arguments presented. These consensus statements were then debated, and agreed upon by workshop participants, and then formulated into the final article in this Special Issue.

This *Value in Health* Special Issue describes the arguments presented during the consensus development workshop on the QALY. Articles 2–6, “QALYs: The Basics,” “QALYs: Some Challenges,” “A Different Approach to Health State Valuation,” “Retaining, and Enhancing, the QALY,” and “The Use of QALYs in Clinical and Patient Decision-Making: Issues and Prospects,” present the arguments. Article 7, “Towards a Consensus on the QALY,” defines eight consensus statements developed by the “consensus developing group” and presents a way forward.

The last two articles in this Special Issue, “Editorial: On The Benefits of Modeling Using QALYs for Societal Resource Allocation: The Model Is the Message” and “Editorial: Moving The QALY Forward or Just Stuck in Traffic?” are provided by two guest editors, selected by the *Value in Health* Editor-in-Chief to critique this Special Issue. The appointment of guest editors is an important component of *Value in Health*’s policy on Special Issues and is intended to ensure that a broad range of views is expressed. The guest editors did not participate in the QALY

Consensus Development Workshop, so their reactions are based on the written articles alone.

Garrison is broadly welcoming of this initiative and considers that the articles in this Special Issue are a fair and useful representation of the current debate about health state utilities and preferences. Nevertheless, he also points out that several important issues were not discussed, or not resolved satisfactorily. He also points out that the relevance of QALYs for resource allocation may depend crucially on the values underlying the organization of the health-care system in a given country. For example, QALYs may have a different relevance in a country like the United Kingdom, with a government-funded national health service, than they may have in a country like the United States.

On the other hand, Reed Johnson is much more frustrated with the lack of progress with research within the QALY paradigm and feels that more progress would be made if we looked for alternatives. These could be alternative methods for elucidating preferences, such as contingent valuation and discrete choice experiments, or different decision-making approaches, such as that currently being proposed by the Institute for Quality and Efficiency in Health Care in Germany. We support further research into, and experimentation with, alternatives to the QALY. As these approaches become more often used in practice, their difficulties will become better known and, hopefully, resolved. Nevertheless, the focus on this workshop was on experience with the QALY. Perhaps, similar workshops, discussing experiences with using alternative measures, will be possible in the future.

We hope that this *Value in Health* Special Issue on “Moving the QALY Forward: Building a Pragmatic Road” is both informative and serves as a platform for future debate and consensus development on the QALY.

Background

The QALY is a widely used measure of both quantity and quality of life. Its benefit for health-care evaluation is that it is applicable to “all individuals and all diseases” and can therefore be used to compare interventions across diseases and programs [2]. It is thus particularly important to outcomes researchers as they attempt to evaluate the efficacy and cost of various health-care interventions and to health-care decision-makers as they weigh implementation or purchase of health-care technologies, including diagnostics, devices, and medications. The QALY is particularly useful because it enables comparisons across diseases, populations, and programs. Some health-care systems, such as the National Institute for Health and Clinical Excellence (NICE) of the British National Health Service (NHS) use QALYs to determine health-care priorities [3]. QALYs are also used in population health, to measure and compare the health of a community.

Address correspondence to: Marilyn Dix Smith, ISPOR, 3100 Princeton Pike, 3E, Lawrenceville, NJ 08648, USA. E-mail: mdsmith@ispor.org
10.1111/j.1524-4733.2009.00514.x

There are several elements to the estimation of QALYs, many of which raise methodological issues. The main element is the assignment of preference values, sometimes known as “utilities” to different health states. To accomplish this, health states need to be classified and described to a respondent, who then values them relative to one another, or an anchor point such as death. Several classification systems exist and these are particularly important in the case of the generic health utility measures, such as the EQ-5D or the Health Utilities Index (HUI).

There are also alternative approaches for eliciting the health state preference values, such as the standard gamble, the time trade-off, and the person trade-off. Each approach has its own underlying rationale and, in some cases, set of axioms. Another key issue is that of “whose values” should be used as the source of values, the candidates including patients, decision-makers, and the general public. The weights or value “tariffs” of all the generic measures are based on surveys of the general public, although the EQ-5D also includes an assessment of self-rated health, which would be from patients when the measure is used in the context of a clinical study.

A second element in the estimation of QALYs is the construction of the profile of health state values over time. In estimating the QALYs gained from a health-care intervention, the analyst calculates the area under the curve, that is, the difference between the profiles obtained for the treatment of interest and its comparator. The simplest, and most frequently used, approach for estimating the profile is to multiply the time spent in each health state over time with its corresponding health state value. Nevertheless, some researchers argue that it is better to estimate the profile holistically, because the value of a given health state is not independent of the time spent in it, or the order in which it is experienced.

A third major element in estimating the total QALYs gained from a given health-care intervention is the aggregation of QALYs across all the recipients. The standard approach is to treat all QALYs gained as being equal, no matter to whom they accrue. Nevertheless, this approach is questioned by some who argue that QALYs may have a different weight depending on, for example, the individual’s initial health state, or the state to which their health is raised.

As the definition, methods, alternatives, and evolution of the QALY are debated, one needs to understand how this tool is being used in health technology assessment and reimbursement decisions. The ultimate use of the tool should be considered in each of the previously described development steps of the QALY to assure continuous improvement in the quality of decisions being made using this tool. The perspective of the end user plays a critical role. In the United Kingdom, a country with a single-payer health system and a clear societal perspective in technology assessment, the QALY seems particularly well utilized in decision-making. This has been demonstrated by the health economist experts who participate on the NICE Review Board for the NHS. The cost and benefit of a technology is summarized in a cost per QALY measure, and standards have been set to create a threshold (generally £30,000 per QALY) of what is acceptable to society. On the other hand, because of a decentralized, fragmented health-care system in the United States, the payers often have different perspectives in making health-care decisions, and these perspectives may or may not include the concept of quality of life. The QALY may be highly relevant to a single payer such as Medicare; however, the drug benefit for the elderly is delivered through an uncoordinated private payer system [4]. There is also a difference in how the QALY is used, for example, as a measure of efficacy for registration or as a utility measure in a cost-effectiveness

ratio for reimbursement. Across the various perspectives and potential uses of the QALY, consistency is a key objective. If alternative definitions, methods, and approaches are used to accommodate differences in perspective, the complexity will inhibit the practical application of the QALY in making better decisions on the allocation of scarce resources in health care.

Thus, despite its benefits and wide use, the methodology and results of attempts to estimate QALYs have been the subject of debate in the health-care field for some time. Outcomes researchers have discussed whether the preference weights required by the measure should come from patients or the community, as preferences between these two groups can differ [2,5]. In addition, different measures used to calculate QALYs, such as the HUI and the EQ-5D, may yield results that are not directly comparable [2,6], and thus not optimal for truly informed decision-making. Finally, decision-makers themselves question whether the QALY can be meaningful in all instances: some feel that the benefits of certain medications that improve the quality of life, such as those for arthritis, cannot be adequately measured by life-years survived. Moreover, some fear that models using QALYs might result in interventions that enhance the quality of life, such as erectile dysfunction medications, appearing a better use of resources than interventions that might extend or save life, such as dialysis or bypass surgery [7].

Perhaps more importantly, the fundamental assumptions behind the QALY have begun to be questioned. Daniel Kahneman, the Eugene Higgins Professor of Psychology and Professor of Public Affairs at the Woodrow Wilson School, Princeton University, and a 2002 winner of the Nobel Prize in Economic Sciences, has observed that “QALYs obtained from patients and the QALYs obtained from the public are going to be radically different. This is because the public and the patients have different reference points” [8]. In a 2005 plenary presentation at the ISPOR 10th Annual International Meeting [9], Kahneman noted that the experience of patients is often quite different from outside observers expect it to be. Studies show that paraplegics, for example, suffer a tremendous negative impact when the condition first occurs, but that the impact moderates over time, which is not expected by outside observers, asked to estimate the effect of such a condition. Paraplegics integrate their condition with other life events and focus less on their condition as time goes by. Such discrepancies highlight the need for a measure that takes patient’s experience into account. Professor Kahneman outlined a paradigm shift between decision utility, which underlies the QALY and is based on inference of choice, and behavioral economics (experienced utility), which is a relatively new field interested in developing optimal methods of measuring patients’ experience [10].

Among outcomes researchers, the interest in defining and using the most meaningful paradigm has stimulated an interest in reviewing the underlying tenets of the field. George Torrance, emeritus professor at McMaster University, has argued that, because the measures used were generated from diverse fields such as economics and decision sciences and then applied to health care, the results may not be fully comparable unless the underlying axioms are specified adequately [11].

In a 2006 Issue Panel debate at the 11th Annual ISPOR International Meeting, Dennis Fryback argued that outcomes researchers, health-care professionals, and decision-makers concerned with questioning the QALY, however, are counterbalanced by another group who argues that the QALY serves a function in health care analogous to the Dow Jones Industrial Average in finance: a useful, if not comprehensive, benchmark understood by all participants and possessing a long history [5]. In addition, some health-care researchers contend that the lack of

a simple better measure as an alternative makes the QALY an indispensable tool. Finally, there is a cluster of questions centering not so much on the QALY per se but around issues of its augmentation or replacement. Many outcomes researchers focus on a simple question: if not the QALY, then what? What would—or could—replace the QALY?

During the 2006 Issues Panel debate, Professor Kahneman suggested that researchers think of what an ideal measure might be, and suggested research into potential tools utilizing the discipline of behavioral economics. Professor Fryback suggested that research and decision-making, in addition to pragmatically using the QALY, could benefit the development of a broader set of data and methods. As noted above, he itemized several additional methods as a starting point: 1) better use of public deliberative processes for valuing health descriptive systems; 2) the collection of data about, exploration of, and integration of longitudinal observations of the health experience of people into the community's deliberative process; and 3) more routine collection and use of existing health-related indexes.

ISPOR Consensus Development Workshop on Moving the QALY Forward

Given the QALY's overall importance to the measurement of health care, its modification, enhancement, or replacement would have significant potential to change health-care methodology and stimulate research going forward. Therefore, ISPOR convened a 2-day invitational consensus development workshop on "Moving the QALY Forward," November 7–8, 2007, Philadelphia, PA, USA with thought leaders in the field with the goal to advance the knowledge in this field, to clarify or transform concepts and methods that drive the field, and to define a research agenda.

The tenets and issues presented and debated at this workshop were as follows.

QALYs: The Basics

The definition of a QALY as a weighted life-year and the taxonomy of methods for deriving those weights were presented and debated.

QALYs: Some Challenges

The development of alternate or preferred definitions of the QALY, defining the metric of the QALY, listing the assumptions, and defining the validity of these assumptions were also presented and debated.

A Different Approach to Health State Valuation

Because experience-based utility may complicate the QALY, a proposal to establish citizen's juries for health state valuation was presented. In addition, the importance of including efficiency, distributive, and procedural justice in the health-care decision-making process was presented.

Retaining, and Enhancing, the QALY

The need to 1) to take stock of current data sets; 2) to outline ways of integrating new work with existing data; and 3) to determine the strengths and weaknesses of existing methods were presented and debated.

The Use of QALYs in Clinical and Patient Decision-Making: Issues and Prospects

The use of measures, such as the QALY, relate to profound social decisions. An improvement in health outcomes might not be the

only reason to use the QALY: other reasons are overall improvement of societal welfare or as an indicator of society's care and compassion. The need to 1) identify how to optimize public deliberative processes for valuing health; 2) discuss barriers in decision-makers' adoption and use of the QALY, with a view toward strategizing improvement of the measure or improvement of decision-makers' understanding; and 3) formulate what the user wants and develop a guide to decision-maker identification of issues for outcomes researchers were presented and discussed.

ISPOR Consensus Development Workshop on Moving the QALY Forward Attendees

The following thought leaders participated in the development of the five group presentations: Diana Brixner, BSPharm, PhD, University of Utah, Salt Lake City, UT, USA; Karl Claxton, MSc, PhD (participated in teleconferences but unable to attend consensus workshop), University of York, York, UK; Norman Daniels, PhD, Harvard University School of Public Health, Boston, MA, USA; Paul Dolan, DPhil, University of Sheffield, Sheffield, UK; Michael Drummond, DPhil, University of York, UK; Dennis G. Fryback, PhD, University of Wisconsin-Madison, Madison, WI, USA; Marthe Gold, MD, MPH, City University of New York Medical School, New York, NY, USA; Daniel Kahneman, PhD, Princeton University, Princeton, NJ, USA; Mark Scott Kamlet, PhD, Carnegie Mellon University, Pittsburgh, PA, USA; Paul Kind, University of York, York, UK; Jennifer Elston Lafata, PhD, Henry Ford Health System, Detroit, MI, USA; Joseph Lipscomb, PhD, Rollins School of Public Health, Atlanta, GA; Karl Matuszewski, MS, PharmD, University HealthSystem Consortium, Oak Brook, IL, USA; Alistair McGuire, PhD, London School of Economics, London, UK; Erik Nord, PhD, Norwegian Institute of Public Health, Oslo, Norway; Dennis Raisch BSPharm, PhD, University of New Mexico Health Sciences Center, Albuquerque, NM, USA; Dennis Revicki, PhD, United BioSource Corp., Bethesda, MD, USA; George Torrance, PhD, McMaster University, Toronto, Canada; Milton C. Weinstein, PhD, Harvard University School of Public Health, Boston, MA, USA.

The following attendees also participated in the discussions during the 2-day workshop: Martin L Brown, PhD, National Cancer Institute, Bethesda, MD, USA; Steven Clouser, PhD, National Cancer Institute, Bethesda, MD, USA; William Lawrence, MD, DrPH, Agency for Healthcare Research and Quality, Rockville, MD, USA; Andreas Maetzel, PhD, MD, MSc, Amgen (Europe) GmbH, Zug, Switzerland; Bryce Reeve, PhD, National Cancer Institute, Bethesda, MD, USA; James A. Shuntinga, PhD, Office of the Director, National Institutes of Health, Bethesda, MD, USA.

We thank Rita Williams, MA, Princeton, NJ, USA for her contribution in developing the grant for this consensus development workshop.

Source of financial support: Funding for the ISPOR "Building a Pragmatic Road: Moving the QALY Forward" Consensus Development Workshop was made possible in part by grant 1R13 HS016841-01 from the Agency for Healthcare Research and Quality. The views expressed in written conference materials or publications and by speakers and moderators do not necessarily reflect the official policies of the Department of Health and Human Services; nor does mention of trade names, commercial practices, or organizations imply endorsement by the US government. Funding for this *Value in Health* Special Issue, "Moving the QALY Forward: Building a Pragmatic Road" was made possible in part by Contract No. HHSN261200800148P from the National Cancer Institute.

Marilyn Dix Smith, Michael Drummond, and Diane Brixner have no conflicts to declare.

References

- 1 Tuckman B. Developmental sequence in small groups. *Psychol Bull* 1965;63:384–99.
- 2 Berger MC, Binglefors K, Hedblom EC, et al. Health Care Cost, Quality, and Outcomes: ISPOR Book of Terms. Lawrenceville, NJ: International Society for Pharmacoeconomics and Outcomes Research, 2003.
- 3 Dawson D, Gravelle H, O'Mahony M, et al. Developing new approaches to measuring NHS outputs and productivity. Summary Report. CHE Research Paper 6. York, UK: Centre for Health Economics, The University of York.
- 4 Brixner D. Reflections on the quality of QALY. *ISPOR CONNECTIONS* 2007;13:3.
- 5 Williams R. Will the QALY survive? *ISPOR CONNECTIONS* 2007;13:13–14.
- 6 Kind P. Incorporating values into healthcare decisions: questioning the Q in QALYs. Slides presented at: First Plenary Session, “Patient Values and Societal Values: How to Measure and Compare Outcomes and What Considerations for the Different Audiences” Presented at: ISPOR 8th Annual European Congress, Nov. 6–8, 2005, Florence, Italy.
- 7 McGregor M. What decision-makers want and what they have been getting. *Value Health* 2006;9:179–80.
- 8 Determinants of health economic decisions in actual practice: the role of behavioral economics. *Value Health* 2006;2:65–7.
- 9 Kahneman D. Determinants of health economic decisions in actual practice: the role of behavioral economics. ISPOR 10th Annual International Meeting, May 15–18, 2005, Marriott Wardman Park Hotel, Washington, DC. Available from: <http://www.ispor.org/meetings/washington0505/051605.asp> [Accessed September 15, 2008].
- 10 Dolan P, Kahneman D. Interpretations of utility and their implications for the valuation of health. *Econ J* 2008;118:215–34.
- 11 Torrance GW. Utility measurement in health care: the things I never got to. *Pharmacoeconomics* 2006;24:1069–78.