



Cost-per-QALY in the US and Britain: Damned if you Do and Damned if you Don't

18th Annual Lecture

Professor Milton Weinstein

15 November 2011
Royal College of Physicians
London

**Cost-per-QALY in the US and Britain:
Damned if You Do and Damned if You Don't
18th Annual Lecture**

**Professor Milton Weinstein
Henry J. Kaiser
Professor of Health Policy and Management,
Harvard School of Public Health, USA**

**15 November 2011
Royal College of Physicians
London**

About the Office of Health Economics

Founded in 1962, the OHE's terms of reference are to:

- commission and undertake research on the economics of health and health care
- collect and analyse health and health care data for the UK and other countries
- disseminate the results of this work and stimulate discussion of them and their policy implications.

The OHE's work is supported by research grants and consultancy revenues from a wide range of UK and international sources.

The views expressed in this publication are those of the author and do not necessarily represent those of the OHE.

Office of Health Economics
Southside, 7th Floor
105 Victoria Street
London SW1E 6QT
United Kingdom
www.ohe.org

©2011
All rights reserved
Printed in the United Kingdom
ISBN 978-1-899040-59-9

Content

1. Introduction.....	3
2. Attitudes in the US and arguments against CEA.....	3
2.1. There is no relation between health care expenditures and health outcomes across hospitals and geographic areas.....	4
2.2. If we eliminate waste, there will be enough money to pay for all useful health care	7
2.3. If we stop overpaying doctors, there will be enough money to pay for all useful health care	7
2.4. If we stop overpaying drug companies, there will be enough money to pay for all useful health care	8
2.5. If we do more prevention, there will be enough money to pay for all useful health care	8
2.6. If we do more comparative effectiveness research, we can identify useless health care, so there will be enough money to pay for all useful health care	8
2.7. QALYs discriminate against the disabled, the elderly and/or children ...	8
2.8. If patients are more involved in decision making, fewer resources will be spent, and patients will get what they want.....	9
2.9. The market will determine the right level of spending on health care, and the right allocation of resources	9
2.10. QALYs don't reflect everything that people care about in healthcare	9
3. Economic evaluation in the UK: my way or the highway?.....	10
3.1 UK guidelines are strict.....	11
4. Who has it right?.....	14
References	15

About the author

Professor Weinstein is the Henry J. Kaiser Professor of Health Policy and Management, Harvard University T.H. Chan School of Public Health; Professor of Medicine at the Harvard Medical School; and Director of the Program on Economic Evaluation of Medical Technology. He is best known for his research on cost-effectiveness of medical practices and for developing methods of economic evaluation and decision analysis in health care.

He is the author of four books and more than 300 papers in peer-reviewed medical, public health, and economics journals. An elected member of the National Academy of Medicine, he is a recipient of the Award for Career Achievement from the Society for Medical Decision Making and the Avedis Donabedian Lifetime Achievement Award from the International Society for Pharmacoeconomics and Outcomes Research.

Professor Weinstein received his AB and AM in Applied Mathematics (1970), his MPP (1972) and his PhD in Public Policy (1973) from Harvard University.

1. Introduction

The purpose of these remarks is not to debate the technical issues of methodology for cost-effectiveness analysis or decision modelling. Instead, the intent is to comment on the marked differences in attitudes in the US and the UK about how cost-effectiveness analysis (CEA) fits into the health care system, particularly CEA based on cost per quality-adjusted life year (QALY). While CEA is an uncommon practice in the US, in the UK there exist very precise guidelines for conducting the analysis.

Important differences in the health care systems of the two countries affect how and how much CEA affects decisions. Health services in England are largely free at the point of use. The National Health Service (NHS), created in 1948, provides preventive medicine, primary care and hospital services. UK residents have the right to use NHS health care without charge, apart from some co-payments required for outpatient prescription and dentistry services (Harrison, 2013; Thorlby and Arora, 2015).

The US system is more diverse. Its national programmes are mostly limited to Medicare, for those aged 65 and older or permanently disabled; Medicaid, which covers the indigent and is partly funded by the states and administered by them; and programmes that cover military veterans and federal government employees (the military, Congress, etc.). In 2011, Medicare accounted for 21 per cent of total national health expenditure and Medicaid for 15 per cent (CMS, 2013). The main role of the private insurance companies in the US health system is evident. In 2014 only 17.7 per cent of the population aged between 18 and 64 years had a public health plan coverage, while 67.3 per cent had a private one (Ward et al., 2015), mostly through employers. Private insurers, of course, may impose their own CEA rules on services and many do, as do some states under their programmes. The 2010 Patient Protection and Affordable Care Act eventually may produce important changes, but probably not in the immediate future.

2. Attitudes in the US and arguments against CEA

The US has not yet embraced cost-effectiveness analysis in government programmes on a national scale – far from it. In fact, the 2010 Patient Protection and Affordable Care Act actually prohibits the federal government from basing policy on QALYs:

The Patient-Centered Outcomes Research Institute established under section 1181(b)(1) shall not develop or employ a dollars-per-quality adjusted life year (or similar measure that discounts the value of a life because of an individual's disability) as a threshold to establish what type of health care is cost effective or recommended. The Secretary [of Health and Human Services] shall not utilize such an adjusted life year (or such a similar measure) as a threshold to determine coverage, reimbursement, or incentive programs under title XVIII [Medicare]. (Enacted by the Senate and House of Representatives of the USA, 2010)

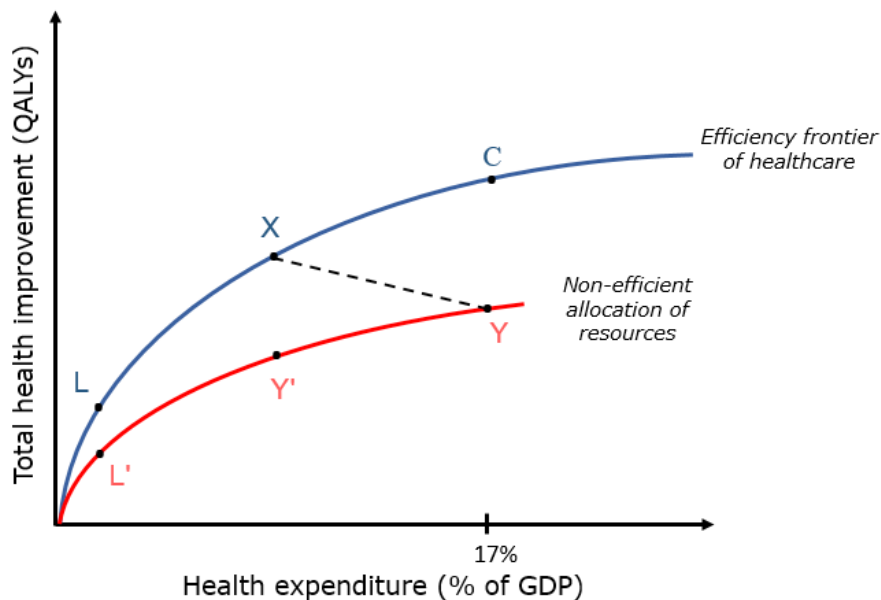
Different arguments against CEA are responsible for the negative opinion that US policy makers have about this tool. In the rest of this section some of the most common reasons used to justify the exclusion of cost-effectiveness from the US decision-making process are briefly analysed. Some of them have some merit and others have none, but they have all been used.

2.1. There is no relation between health care expenditures and health outcomes across hospitals and geographic areas

Certain parts of the literature suggests that the relationship between overall expenditure and outcomes across geographic areas or across hospitals tends to be fuzzy or non-existent. In fact, there are studies that indicate a negative relationship across geographic areas (Baicker and Chandra, 2004).

This somewhat counterintuitive result has been challenged by Weinstein and Skinner (2010), whose hypothesis links the need to consider cost-effectiveness analysis to the negative or weak relationship between expenditures and outcomes. The argument is illustrated in Figure 1, which shows the “efficiency frontier of health care”. The efficiency frontier of health care represents the situation in which there is a limited budget to spend, and this limited budget is allocated first to the most cost-effective intervention (a point such as L, Figure 1) and then to the next most cost-effective (moving to a point such as X, Figure 1). The movement along the curve continues until the total budget has been spent (for example, at a point such as C, Figure 1). If the process of allocation of resources is financing the most efficient treatments, the end point is such that the highest possible health improvements are funded. The health improvements could be measured, for instance, in terms of QALYs on the vertical axis in Figure 1.

Figure 1. Levels of efficiency in allocating health care resources



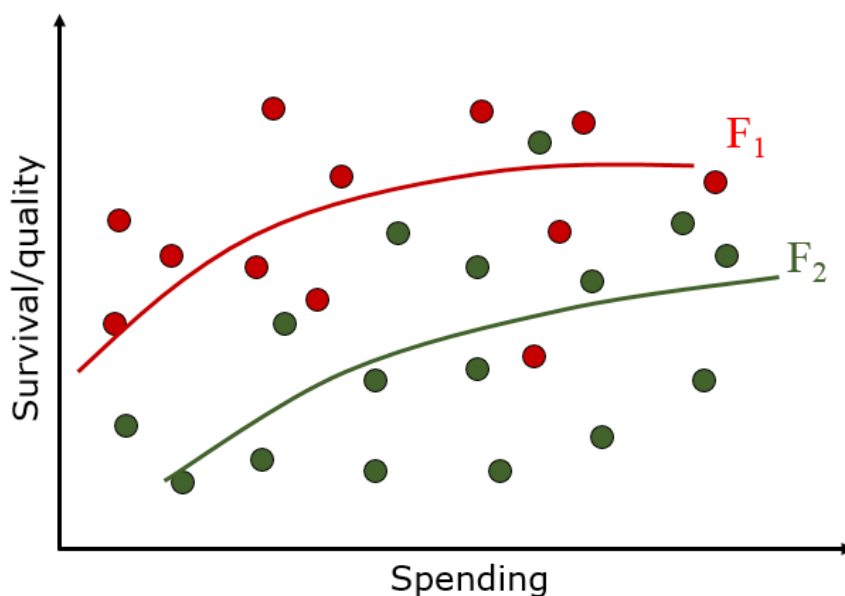
Source: Weinstein and Skinner (2010)

The findings of a negative relationship between overall health expenditures and health outcomes could be explained by the fact that in reality health care expenditures are not used efficiently. There are many situations in which the most cost-effective interventions are underutilised. For instance, in the US 17 per cent of gross domestic product is spent on health care and the percentage is growing continuously. Instead of spending 17 per cent of GDP on health care treatments that could produce a level of improvement such as the one observed at point *C* (Figure 1), the 17 per cent is spent on a selected group of treatments that allow the achievement of a lower level of health improvements such as the one observed in point *Y* (Figure 1). For example, fewer than half of Americans over the age of 50 have ever had a colorectal screening exam, which is generally regarded as a very cost-effective intervention; people do not get their influenza vaccinations or pneumococcal vaccinations as recommended.

We considered the case of two geographic locations: region *A* that is rather efficient and is located at a point such as *X*, and region *B* that spends more but in a less efficient way and is therefore located at a point such as *Y*. In a cross-sectional analysis the relationship between health improvement and overall expenditures would be identified as negative (represented by a dashed line in Figure 1) – however, it could be reflecting differences in the allocation of resources (represented by the different slopes of the curves in Figure 1).

Although region *B* may spend more and get less health improvement, this does not mean that it is possible to cut back on health services in region *B* without compromising health. If the level of expenditure in region *B* is cut and the health care sector does less of what it is already doing, region *B* could end up at a point such as *Y'*, with a lower level of health improvement. Region *B* would be compromising health in order to cut expenditure. For region *B* to reduce expenditures and improve health outcomes simultaneously, it is essential to increase the utilisation of some highly cost-effective interventions and at the same time decrease the utilisation of less cost-effective interventions. Consequently, there exists a need in region *B* for cost-effectiveness analysis.

Figure 2 is a qualitative representation of some quantitative data that Skinner, Staiger and Fisher (2006) analysed relating health care spending in hospitals in the year following an acute myocardial infarction (MI) and one-year survival of patients having acute MIs in those hospitals. The authors divided the hospitals into two groups based on productivity, in terms of the rate at which they adopted highly cost-effective technologies to improve outcomes following acute myocardial infarction (F1 hospitals are more productive than F2 hospitals). When we observe the whole set of points no clear pattern emerges between expenditures and outcomes. However, if the hospitals are divided into two groups, there is in fact an increasing relationship between spending and outcomes as much in the F1 hospitals as in the F2 hospitals.

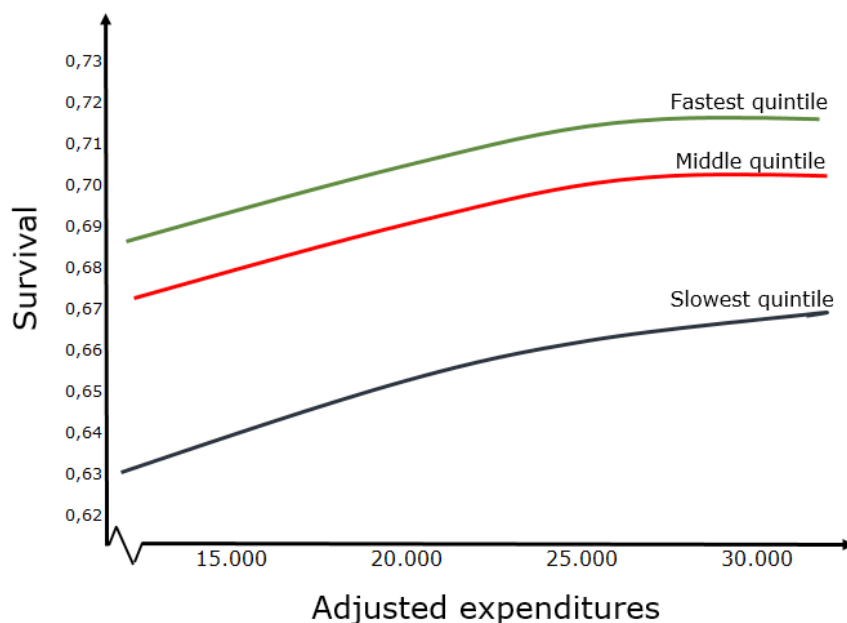
Figure 2. Spending and health outcomes: different “production functions”

Source: Weinstein and Skinner (2010)

In an analysis conducted by Skinner and Staiger (2009), the authors looked at the rate of adoption of three highly cost-effective technologies for acute MI: aspirin, beta-blockers and reperfusion. Currently, these are standard procedures in almost every hospital in the US, but during the time of their adoption (between the 1980s and 1990s), there were important differences in the rate at which each hospital implemented these technologies. Skinner and Staiger (2009) aimed to relate outcome and survival after MI with the following two variables: (1) the hospital level of adoption of the three technologies, and (2) resource expenditure during the next 12 months.

The results suggest that although survival increases in every hospital, the improvement is always higher in the hospitals that were quickest to adopt aspirin, beta-blockers and reperfusion. Considering the rate of adoption of these cost-effective technologies, the regression analysis estimated by Skinner and Staiger (2009) indicates that the fastest-adopting quintile of hospitals have better outcomes than the slowest. Even more surprising is that a positive relationship was found between expenditures and outcomes in all the strata.

Figure 3 is an illustration of Skinner and Staiger's findings. In summary, what the studies of Skinner, Staiger and Fisher (2006) and Skinner and Staiger (2009) suggest is that it is possible for a hospital to cut costs and improve outcomes by adopting cost-effective technologies more rapidly. A hospital can achieve this goal by moving from the slowest quintile in Figure 3 to the fastest or the middle quintile.

Figure 3. AMI survival versus normalised expenditures in US hospitals (1986–2004)

Source: Adapted from Skinner and Staiger (2009)

2.2. If we eliminate waste, there will be enough money to pay for all useful health care

In the US there exists a notion that there is much “waste” in the health sector. However, there is not a clear understanding to which kind of waste people are referring. Certainly in the US there are many incentives to reduce waste in the health sector. For instance, there is competition among hospitals for the business of health care payers, and there is also competition among payers for employers to offer them to their employees. Consequently, it is difficult to understand, with this level of competition, in what manner waste in the health sector is still considered a problem. Even more important, there is evidence that there is less waste in the health system than there was 20 years ago. The fact is that there do not seem to be many opportunities to eliminate much more waste.

2.3. If we stop overpaying doctors, there will be enough money to pay for all useful health care

Another argument suggests that the US health system is paying too much on the supply side; particularly there is a notion that doctors are overpaid. However, the US health system does not pay doctors on average much more or any more than doctors are paid in other countries. The difference is in the organisation of the payment system and incentives. The form in which clinicians are paid in the US is mainly on a fee-for-service basis, so that the more they do the more they receive. Meanwhile, in many other countries clinicians get paid more on a capitation or salary basis.

Even more important, if doctors in the US are paid less, the quality of the health care could be affected since the system may not get enough good-quality doctors or some of the current clinicians might leave the profession. For example, the US does not have as many primary-care doctors per capita as the UK. This is partially explained by the fact that primary-care doctors do not get paid as much as specialists. It does not pay to go to medical school and take a long-term loan to work as a primary-care doctor; it is more profitable to become a surgeon, an anaesthesiologist or a radiologist, or perhaps a sub-specialist in medicine, like a dermatologist.

Therefore a decrease in the salary of clinicians is also not the most viable solution for cutting costs and increasing health outcomes.

2.4. If we stop overpaying drug companies, there will be enough money to pay for all useful health care

In the US there exists the idea that drug companies are being overpaid, and that a decrease in their payments would save enough of the health budget to invest in all needed health care. The fact is that drug companies have to innovate, and a decline in revenues would negatively affect their incentive to do so. Studies suggest that this could be the case. In any case, cutting prices for pharmaceuticals is certainly not the solution to the health care cost problem.

2.5. If we do more prevention, there will be enough money to pay for all useful health care

The topic of prevention was actually used in the political campaign of 2008 by both Mr Obama and Mr McCain. The main argument is that it is only necessary to increase expenditure in disease prevention, and by doing so the health care system would eventually save an important part of its budget, which could be allocated to other needed health care services. This will allow the fulfilment of the demand for health care.

However, even if prevention is beneficial and in some cases cost-effective, the fact is that most prevention does not save money. In a systematic review of the cost-effectiveness studies of preventive services conducted by Cohen, Neumann and Weinstein (2008), the findings suggest that there is essentially no difference between the cost-effectiveness of prevention and preventive services and the cost-effectiveness of treatments. Similarly, they found that at least in the published literature, there is no difference between the likelihood of preventive service being cost-saving compared to the likelihood that a therapeutic intervention will be cost-saving.

2.6. If we do more comparative effectiveness research, we can identify useless health care, so there will be enough money to pay for all useful health care

This argument is based on the idea that by conducting comparative effectiveness analysis of health interventions, it is possible to identify those interventions that are useless. The exclusion of these treatments from the health care services offered will save enough from the health care budget that it will be feasible to pay for everything that it is useful.

The problem is that it is very difficult to prove that an intervention is useless. Randomised trials, if they are feasible, are not intended to prove a negative; they are intended to prove a positive. Therefore, although it could be the case that the treatment cannot be proved to be better than its alternative, it is very hard to show that it is exactly equivalent to the alternative. Moreover, most interventions indeed do not lend themselves to randomised clinical trials and therefore it is necessary to rely on other sources of evidence, and it is very hard to prove beyond reasonable doubt that an intervention is absolutely useless. Consequently, it is difficult to find examples of interventions that have been shown to be completely useless

2.7. QALYs discriminate against the disabled, the elderly and/or children

The QALY is a measure of the disease burden that considers the quality and the quantity of the life lived. This is normally used as a part of the cost-effectiveness

analysis to assess the value for money of a medical intervention.

The group that is opposed to the use of cost-effectiveness analysis has also questioned the capacity of the QALY to reflect the potential benefits of a new treatment. They argue that QALYs are discriminatory against some of the most vulnerable groups in society: the disabled, the elderly and children. The argument is that QALYs discriminate against the elderly because there are fewer life years to gain if you save the life or prevent the death of someone who is elderly. In relation to children, they argue that even though you save more life years you have to discount them at 3 per cent or 3.5 per cent per annum, so either way you are discriminating.

However, the reason for using weights to differentiate health states is that society values life in better health more than in life in less good health. Moreover, QALYs have been never used to discriminate against the disabled. As far as the author is aware, there is no study that suggests the prioritisation of non-disabled people over disabled people.

2.8. If patients are more involved in decision making, fewer resources will be spent, and patients will get what they want

There is some validity in this argument; if patients are more involved in decision making, the health system could save resources since people will feel more empowered to say "I don't want that surgery" or "I don't want that invasive procedure".

On the other hand, it has been noted that in the US people have a high demand for diagnostic tests. They are highly interested in knowing what is actually happening inside their bodies. For example, the demand for CT scans and MRIs is significant. Empowering people to decide which diagnostic tests will be carried out will not necessarily result in fewer health services or a reduction in expenditures.

2.9. The market will determine the right level of spending on health care, and the right allocation of resources

One of the main arguments against the utilisation of cost-effectiveness analysis is that, assuming informed consumers and stable insurance markets, the interaction between market forces will assure the correct level of health care expenditures.

This argument is the core of the differences between the US and the UK. Americans believe in the proper functioning of market forces. The Americans trust these forces to organise the health care system by determining the right level of spending on health care and the right allocation of resources. However, it has been proved that market forces do not work particularly well in health care. There are many reasons for this; asymmetry of information is one of the most important. In effect, the evidence suggests that markets have not worked very well in the US health care system where there has been a continuous cost escalation without much improvement in outcomes

2.10. QALYs don't reflect everything that people care about in health care

This is the argument that could have the most merit, but it is almost never considered in political discourse. QALYs do not reflect everything that people care about in health care. For example, there may be value in some genetic testing that tells people what risks they face as they proceed through life or what risks their child faces. Even if you cannot do anything about it, there is the psychological value of knowing. Another example is caring, which does not necessarily manifest itself in more QALYs but is something that people value, such as going to a doctor and having the doctor actually spend some time talking to the patients. Similarly, other factors such as access to care, equity and reducing disparities in society are valuable for the population but are not reflected in maximising QALYs.

This argument should be kept in mind for those who are enthusiasts about cost-effectiveness analysis because it reflects some values that are held highly as much in the US as in the UK and reveal some of the limitations of cost-effectiveness analysis. In this regard, one of the most important recommendations of the US Panel on Cost-Effectiveness in Health and Medicine was that cost-effectiveness analysis is an aid to decision-making, not a complete procedure for making resource-allocation decisions, because it cannot incorporate all the values relevant to such decisions (Weinstein et al., 1996). Another of the suggestions from the panel is the so-called rule of reason, which refers to the fact that the analysis should not be more complicated than it needs to be in order to inform the decision to which the analysis is addressed. This rule should be applied to all recommendations about what health outcomes and costs should be included and how to measure them.

These recommendations bear in mind that the cost-per-QALY ratio is only one consideration relevant to the optimal use of health care resources, and should not make the analysis more complicated than it has to be. Those of us who support and apply cost-effectiveness models could be more mindful of this than we are. It is necessary to be aware of the limitations and be mindful of the role that this type of analysis has among many other considerations, such as the ethical and the psychological. Among the assiduous supporters of cost-effectiveness analysis is the UK through the institution that is responsible for the assessment of new treatments: NICE (National Institute for Health and Clinical Excellence).¹ In the following section the application of cost-effectiveness analysis during the decision-making process in the UK context (England and Wales) is discussed.

3. Economic evaluation in the UK: my way or the highway?

The US Congress's prohibition of using cost per QALY in federally funded programmes is the opposite of practice in the UK. Cost per QALY is not only used in government decision making in the UK, but is actually mandated as part of the guidance for decisions about coverage and reimbursement under the NHS (NICE, 2013).

In the UK, the cost per QALY of interventions is calculated according to strict guidelines (NICE, 2013); lifetime costs and QALYs usually are estimated by mathematical models. The question is whether the British take their prescribed guidelines for cost-per-QALY modelling too seriously.

Developing recommendations on good research practices in modelling studies was the focus of a multinational task force led by Professor Andrew Briggs from the University of Glasgow (Briggs et al., 2012). The recommendations included these points on the purpose and use of models in health care:

- The purpose of a model is to inform medical decisions and health care resource allocation
- Modellers employ quantitative methods to gain qualitative insight
- The tools of formal analysis are employed to structure the clinical, epidemiological and economic evidence base in service of better clinical practice decisions and public health priorities

What this says is that models are important, but they are a means to an end, not an end in themselves. What it does not say is that cost-per-QALY estimates resulting from models should be used to make medical decisions.

¹NICE was renamed the National Institute for Health and Care Excellence on 1 April 2013.

The great British mathematician George Box said, "All models are wrong, but some are useful" (Box and Draper, 1987). In fact, models often are not only useful, but essential for a number of reasons. Clinical trials, for example, are necessarily limited, time horizons are short and each trial can address only a limited set of options (no more than two or three, when in reality options may be numerous). Take screening programmes, for example, where decisions must be made about what test to use, how often to use it, who should get it, and what to do if the test result is positive.

Clinical trials also always have restricted target populations that may not mirror the real world well. Often, the trial population is selected purposely to provide the high-risk group required to measure numerous outcomes and endpoints. By necessity, in addition, endpoints are quite simple in clinical trials, and, also by necessity, they lack the flexibility to adopt new evidence from other sources, including observational data.

3.1. UK guidelines are strict

The guidelines for doing cost-per-QALY studies in the UK are very strict. Two examples are the use of the EQ-5D and probabilistic sensitivity analysis (PSA).

The EQ-5D is a scale for measuring and valuing health outcomes that was developed by a multinational European Panel, EuroQol (Kind, 1996). It consists of a five-item scale of health, with weights from zero to one based on preference surveys carried out in the community. The UK guidelines deem it the best scale, the one that should be used in all economic evaluations for studies presented to NICE, which makes recommendations on coverage that are binding across the NHS.

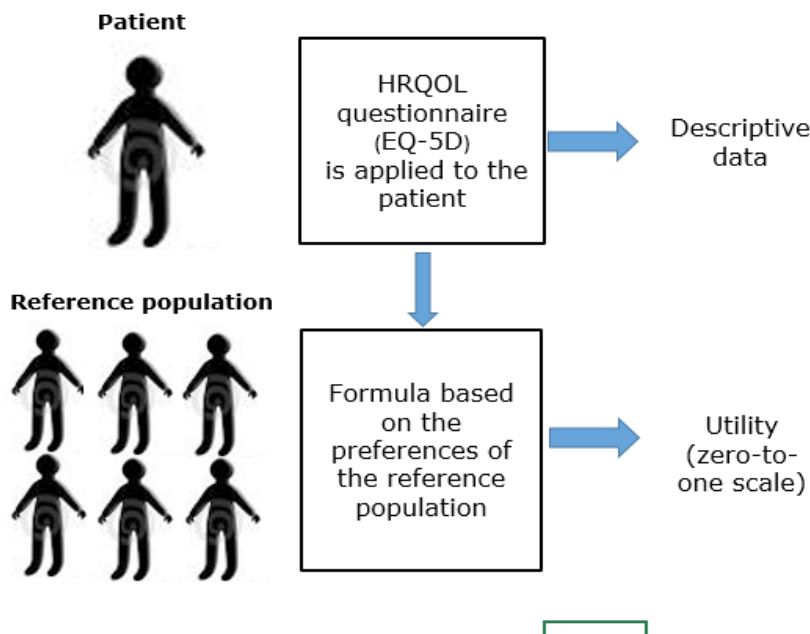
Table 1 explains the elements that are considered in the EQ-5D. It contains five domains of health: (1) mobility, (2) self-care, (3) usual activities, (4) pain or discomfort and (5) anxiety or depression. Each of these elements can exist at three levels in an individual at any given point in time, ranging from no problems to some problems to severe problems.

Table 1. The EQ-5D health scale

	Mobility	Self-care	Usual activities	Pain/ Discom- fort	Anxiety/ Depression
Level 1: no problems	No prob- lems walking about	No problems with self- care	No problems with performing usual activities	No pain or discomfort	Not anxious or depressed
Level 2: some prob- lems	Some problems walking about	Some problems washing or dressing self	Some problems with performing usual activities	Moderate pain or dis- comfort	Moderately anxious or depressed
Level 3: severe prob- lems	Confined to bed	Unable to wash or dress self	Unable to perform usual activities	Extreme pain or dis- comfort	Extremely anxious or depressed

In economic evaluation, an actual or modelled (virtual) individual is placed in one of these cells, as is shown in Figure 4. Each cell has associated with it a weight on a zero-to-one scale that represents the preferences of a reference population. In practice, the weights can be less than zero; some health states can have negative weights that imply that a health state is considered worse than being dead.

Figure 4. The generic-utility approach: patient's health state, community preferences (Hunink et al., 2014)



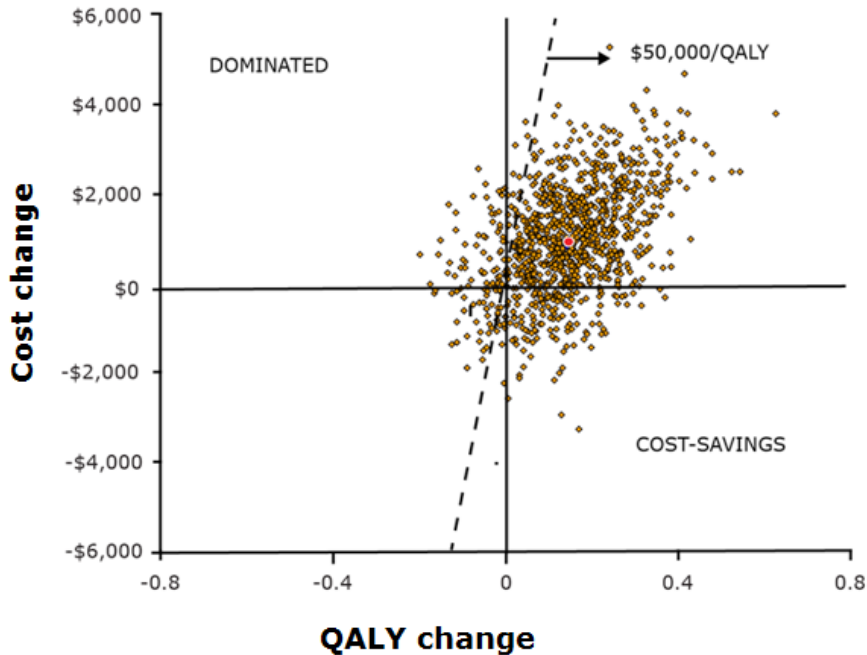
This approach has some “non-fatal” problems that nevertheless raise doubts about whether this is the ideal instrument for measuring health status. For example, the best health state short of perfect health is one where the patient has some problems walking about, perhaps using a cane or maybe with a limp. This health state has a weight of 0.88 in the EQ-5D system, meaning that it is 12% worse than perfect health. When it comes to calculating QALYs, that impairment is equivalent to losing 12% of total life expectancy. Is it realistic to assume that people really believe that this health state is equivalent to living 12% fewer years of life in perfect health? This may limit how useful this measure can be in evaluating relatively mild impairments that might be relieved through treatment.

A second problem is that having only five domains means that some important aspects of health are not captured. Aspects such as fatigue, energy, nausea and sexual function, which are important components of human health, are not reflected in the five domains listed in the EQ-5D.

Regarding probabilistic sensitivity analysis (PSA), it is first worth mentioning the two methodologies normally used to examine the uncertainty around the expected output of models (e.g. cost per QALY). The first methodology is deterministic sensitivity analysis, which is relatively straightforward. This explores changes in values or assumptions that affect cost-effectiveness estimates such as treatment efficacy or costs – for instance, the effect on the cost-effectiveness if the treatment efficacy is changed to the upper 95 per cent confidence limit or lower 95 per cent confidence limit from the trial.

The second methodology normally used is PSA. This answers a more demanding question: what is the probability that an intervention is cost-effective at, say, a £30,000-per-QALY threshold? This method requires specification of probability distributions around all the model parameters: not just efficacy (which is not an issue if determined in a clinical trial), but also the rate of adverse events, costs and utilities. These distributions are calculated simultaneously, producing a distribution for the cost-per-QALY.

Figure 5 shows a scatter plot for a probabilistic sensitivity analysis (Rinfret et al., 2005). Each dot represents a parameter set drawn from the joint probability distribution of the input parameters. Here, the fraction of the dots that are to the right of and below the dashed line suggest that the probability is around 80 to 85 per cent that this particular intervention is cost-effective at the \$50,000-per-QALY level.

Figure 5. Joint distribution of incremental cost and incremental effectiveness

Although the joint distribution of incremental cost and incremental effectiveness in **Figure 5** is a refined way of representing the probabilities, the reliability and accuracy of the probability distributions around parameters can be an issue. Sometimes these probability distributions are nothing more than best guesses, particularly when no trial-based or empirical evidence exists, whether from clinical trials or from observational studies. Despite this, in NICE guidelines, PSA is preferred for the analysis of uncertainty in the cost-effectiveness analysis (NICE, 2013).

A joint task force of the International Society for Outcomes Research (ISPOR) and the Society for Medical Decision Making (SMDM) has developed a set of guidelines – “good research practices” – that, among other things, address concerns about the use of PSA (Caro, Briggs and Kuntz, 2012). These include the following.

1. Uncertainty analysis can be either deterministic or probabilistic and often it is appropriate to report aspects of both types within a single evaluation. There exists the notion that the estimation of the PSA makes the deterministic sensitivity analysis unnecessary. However, compared to PSAs, deterministic analyses are relatively easy to understand, relatively flexible and much more transparent. The range of parameter values to which each parameter was subjected can simply be stated, rather than explaining and justifying a normal distribution, a beta distribution, a chi-squared distribution or a gamma distribution around each of the parameters in a complex model.
2. PSA is most useful when an option exists for collecting additional data through, for example, clinical studies or post-marketing surveillance. PSAs are not particularly useful unless new data can be collected to suggest how a decision may need to be changed. This recommendation is very important and should be followed not only in the US and the UK, but globally.
3. Parameters for probabilistic sensitivity analyses, or parameter ranges

for deterministic sensitivity analyses, should be disclosed and justified. Published research has sometimes omitted important specifics, such as the underlying probability distributions or the data used to create the PSA.

As I mentioned above, the US Panel on Cost Effectiveness in Health and Medicine holds the view that cost-effectiveness analysis, and economic evaluation in general, is only one of many inputs to decision making about health care resource allocation. Decision making should be guided by many considerations and types of data analysis, with models being just one. Moreover, when a model is used, it is crucial that it be “fit for purpose”. For example, different models will be appropriate for decisions meant to be final than for decisions meant to allow use while more evidence is collected. Using PSA to calculate the expected value of information can help determine whether it is worthwhile to do additional studies to guide future decisions about a particular technology, but PSA alone need not serve any particular decision-making purpose.

Other types of analysis that may form the basis for decisions include “cost-consequence” analysis, which assumes that just identifying the pros and cons (the benefits, harms and costs) of a health intervention is enough and that value judgements should be left to decision makers. This disaggregated information may be difficult for decision makers to process cognitively without some help in organising it. Even when QALYs are not accepted as the best measure, it is important to consider ways to structure and aggregate the outcomes that produce a QALY.

Deliberative processes also can aid decision makers. Conversations that include all stakeholders, including the general public, and address issues such as cost-benefit trade-offs and the ethical components of decision making are important in understanding what matters in health care.

4. Who has it right?

Unquestionably, both policy makers and decision makers in health care in the US should pay greater attention to cost-effectiveness. This will have to happen sooner or later. In this regard, it is worth remembering that QALYs and markets are not mutually exclusive. A possibility for the future application of the cost-effectiveness analysis in the US is to include the concept of QALYs as a tool in competition. It is normal to hear advertisements mentioning quality – if the evidence around cost-per-QALY is open for discussion and inspection, there might be the possibility of some competition between health care delivery organisations in terms of the QALYs offered to potential customers. Health care delivery organisations would like to show that they offer more value for money than their competitors. For instance, Medicare could pay different amounts to the providers depending on their effectiveness or cost-effectiveness. Similarly, a patient could have a zero co-payment for something that is either cost-saving or highly cost-effective and a higher co-payment on a treatment that is less cost-effective – so-called value-based insurance.

Given the pluralism in the US health care system, however, nationwide adoption of any one approach to CEA seems a distant prospect. The advantage may be that this plurality will encourage the further refinement of QALYs, input measures and even other models.

In the case of the UK, measures have been taken to include a broader number of factors, apart from the cost-per-QALY ratio, in the decision-making process. For instance, there is some deliberation and public discussion.

If forced to choose, I would say Britain has it closer to right, but Britain may go a little too far in its tendency to virtually enshrine cost per QALY as the basis for decisions.

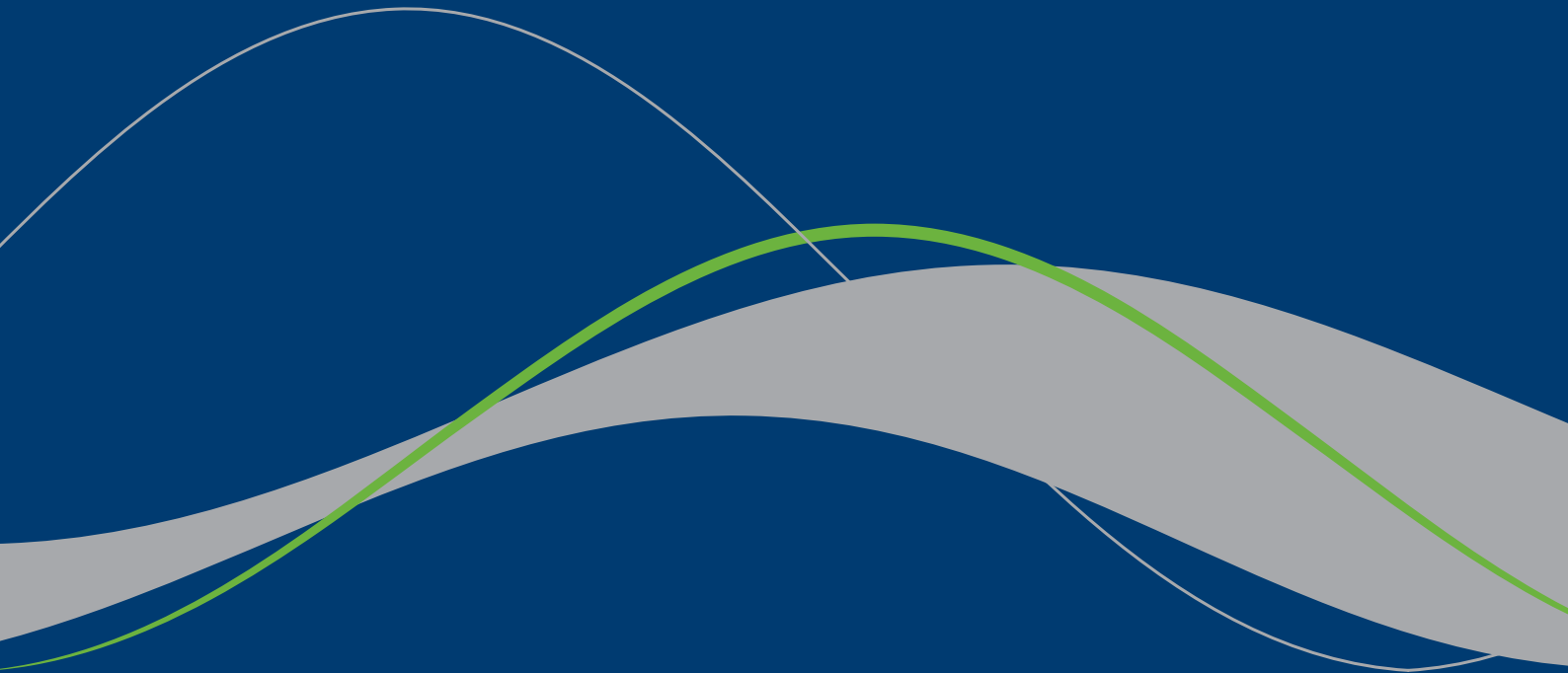
References

- Baicker, K. and Chandra, A., 2004. Medicare spending, the physician workforce, and beneficiaries' quality of care. *Health Affairs (Millwood)*, W4, pp.184–197.
- Box, G. and Draper, N., 1987. *Empirical model-building and response surfaces*. New York: John Wiley & Sons, Inc.
- Briggs, A., Weinstein, M., Fenwick, E., Karnon, J., Sculpher, M. and Paltiel, A., 2012. Model parameter estimation and uncertainty: a report of the ISPOR-SMDM modeling good research practices Task Force-6. *Value in Health*, 15(1), pp.835–842.
- Caro, J., Briggs, A.S.U. and Kuntz, K., 2012. ISPOR Task Force reports: modeling good research practices – overview: a report of the ISPOR-SMDM modelling good research Practices Task Force-1. *Value in Health*, 15(1), pp.796–803.
- CMS (Center for Medicare and Medical Services), 2013. NHS fact sheet. [Online] Available at: <http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NHE-Fact-Sheet.html> [Accessed 16 June 2013].
- Cohen, J., Neumann, P. and Weinstein, M., 2008. Does preventive care save money? Health economics and the presidential candidates. *New England Journal of Medicine*, February, 358(7), pp.661–663.
- Enacted by the Senate and House of Representatives of the USA, MAR. 23, 2010. Public Law 111–148: The Patient Protection and Affordable Care Act. Sec. 1182 (e). USA.
- Harrison, A., 2013. The English health care system. In: S. Thomson, R. Osborn, D. Squires and M. Jun, eds. *International Profiles of Health Care Systems*. New York: The Commonwealth Fund, pp.57–65.
- Hunink, M.G.M., Weinstein, M.C., Wittenberg, E., Drummond, M.F., Pliskin, J.S., Wong, J.B. and Glasziou, P.P., 2014. *Decision making in health and medicine: integrating evidence and values*. Cambridge: Cambridge University Press.
- Kind, P., 1996. The EuroQol instrument: an index of health-related quality of life. In: B. Spilker, ed. *Quality of life and pharmacoeconomics in clinical trials*, 2nd edn. Philadelphia: Lippincott-Raven, pp.191–201.
- NICE (The National Institute for Health and Care Excellence), 2013. *Guide to the methods of technology*. London: NICE.
- Rinfret, S., Cohen, D.J., Lamas, G.A., Fleischmann, K.E., Weinstein, M.C., Orav, J., Schron, E., Lee, K.L. and Goldman, L., 2005. Cost-effectiveness of dual-chamber pacing as compared with ventricular pacing for sinus-node dysfunction. *Circulation*, 111, pp.165–172.
- Skinner, J. and Staiger, D., 2009. *Technology diffusion and productivity growth in health care*. NBER Working Paper No. 14865. Cambridge: National Bureau of Economic Research.
- Skinner, J., Staiger, D. and Fisher, E., 2006. Is technological change in medicine always worth it? The case of acute myocardial infarction. *Health Affairs (Millwood)*, 25(2) pp.w34–w47.
- Thorlby, R. and Arora, S., 2015. The English health care system, 2014. In: E. Mossialos, M. Wenzl, R. Osborn & C. Anderson, eds. *International Profiles of Health Care Systems, 2014*. New York: The Commonwealth Fund.
- Ward, B., Clarke, T.C., Freeman, G. and Schiller, J.S., 2015. Early release of selected estimates based on data from the 2014 national health interview

survey. Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics.

Weinstein, M., Siegel, J., Gold, M., Kamlet, M. and Russell, L., 1996. Recommendations of the panel on cost-effectiveness in health and medicine. *Journal of the American Medical Association*, 276(15), pp.1253–1258.

Weinstein, M. and Skinner, J., 2010. Comparative effectiveness and health care spending: implications for reform. *New England Journal of Medicine*, 362(5), pp.460–465.



Office of
Health
Economics

Research & Consulting