

Commentary

QJM

The cost of a QALY

R. KIRKDALE¹, J. KRELL², C. O'HANLON BROWN², M. TUTHILL² and J. WAXMAN²

From the ¹Medical Student, Imperial College and ²Oncology, Imperial College, London

Address correspondence to J. Waxman, Department of Surgery and Cancer, Division of Cancer, Imperial College London, Room 1014, Garry Weston Centre, Hammersmith Hospital, Du Cane Road, London W12 0NN. email: j.waxman@imperial.ac.uk

Summary

Background: Current regulation of drug approvals has caused considerable controversy as entrusted to the National Institute of Clinical Excellence, and has led to a lack of availability of modern medicines on the basis of calculations made of 'value'.

Aim: We have examined the assessment tool used by National Institute of Clinical Excellence (NICE) to establish the cost of drugs in order to assess whether it is a reasonable and objective evaluation methodology.

Design: A review of the methods of analysis.

Methods: An objective assessment of the value of the Quality Adjusted Life Year (QALY).

Results: We conclude that current methods used by NICE to assess drug costs are arbitrary, subjective and fail to reflect the true costs for patients, which are grossly overestimated.

Conclusion: NICE needs to look again at the evaluation methods for calculating drug costs, and change their methodology from a subjective to an objective measure of true cost.

Quality of life

The Secretary of State for Health set up the National Institute of Clinical Excellence (NICE) in 1999 with the initial brief for the Institute to assess the effectiveness of new treatments as they were developed and brought to market. Somehow the original prescription for NICE was lost; and the brief expanded, with cost to be evaluated in addition to efficacy. The instrument to be used to assess value for money was the Quality Adjusted Life Year (QALY), a loosely defined and subjective costing tool.

In 1999 the QALY had not been established as an objective measure. The use of the QALY in NICE assessments was rejected by the civil servants, clinicians and economists advising the Secretary of State (personal correspondence), but their advice was overridden and the QALY set up as the pivotal instrument for the evaluation of the advantages and disadvantages of a new treatment.

There are many claims and innumerable needs for resources to pay for health, and it is quite clear that the country's purse is not deep enough to pay for every request and fund every need. It is absolutely clear that our resources are not infinite, even though our requirements might be, and nowhere is the conflict between needs and available funds apparently more poignant than in cancer care. We are at a point in cancer drug development where a whole new generation of treatments is becoming available for use. These treatments exploit the molecular characteristics of tumours that cancer research has elucidated. The new generation of treatments is categorized as the 'ibs and abs'. The 'ibs' target cancer cell tyrosine kinase pathways, while the 'abs' are antibodies directed against cell surface antigens specific to individual cancers. These agents are clearly not cheap, and their cost is calculated on the basis of trying to recoup research costs,

which are in the order of £1 billion for each new cancer drug, and make a profit to ensure shareholders are happy and Chief Executives are kept in post. This profit is also required to support research to source the next generation of profit making treatments. However, it is now felt amongst many that NICE functions as a brake on Pharma profits, and not solely a tool to evaluate efficiency.

NICE drugs, the UK and the world

How are drug costs regulated in other areas of the world? In many countries costs are negotiated directly between government and drug companies, or calculated on the basis of an agreed cost principle formula. But, when the relationship between Pharma and Government is considered, the UK is not Canada, nor is it part of Europe. In the UK decisions are made on the basis of the QALY 'formula'. But, there is no unified policy for the UK and decisions on cost effectiveness, based on the QALY, are worked out by individual institutes in UK. NICE has become notorious recently for the length of time that it takes to make a judgment, the ease with which it seems to change its mind about judgments, and the disagreement that the community of doctors and patients have with NICE as to the virtue of their cost calculation. For example, the recent renal cell cancer drug approvals took nearly 3 years to process from publication of definitive trial results in the *New England Journal of Medicine*, with NICE changing its decision on approval three times.¹ This is in contrast with Food and Drug Administration approval, which took <1 month.

The QALY

The NICE cost calculation is based on a QALY, which is a measure of disease burden, encompassing both the quality and the quantity of life lived. The QALY establishes and defines the cost of a new treatment or a health care intervention. By this means the QALY can be applied to provide a value for these treatments and interventions that can be used for comparison between new and established treatments. The QALY has become a well worn key that allows us to turn the lock of a cost-benefit-analysis of medical treatments.

The QALY impacts upon the lives of patients and doctors because of its critical involvement in NICE assessments. These appraisal processes conclude with a recommendation from NICE, upon which healthcare professionals and the National Health Service (NHS) are obliged to act.² NICE, de facto, decides what drugs are available for patients, and

not doctors. But what actually is the QALY, and what has the QALY replaced? In the process of drug testing, safety and efficacy data has always been paramount. If a new treatment was safe and more effective than an older treatment, improving upon response and its duration, if it wasn't toxic and was not too costly, it would be recommended to patients. This description of activity has been replaced by the QALY because the QALY takes into account quality of life in addition to response rates and absolute survival, seemingly objectifying judgments that were subjective. However, as we will show, the description of quality of life remains subjective and difficult to standardize, and puts a seemingly scientific gloss on what remains a subjective and variable formula.

Quality of life

So, what is quality of life, and can it be standardized? As might be obvious, the definition of quality of life is highly subjective, as individuals have different concepts of what represents life quality that varies from day to day, and changes in the context of good and bad health. The QALY was developed to provide an objective means to overcome this subjectivity and provide a set of standardized values for quality of life. However, in our view, it does no such thing and is therefore not fit for purpose. To establish a value for calculating a QALY, quality of life is rated between 0 and 1. On this scale 1 represents 'best possible health' and 0 'worst possible health' (a euphemism for death). There is more than one approach to generating a quality of life valuation, with some assessment tools using a scale that includes negative values, reflecting the concept of a health state worse than death. In some scales, conditions of health are given a decimal value where, for example, being bedbound might be scored as 0.5.

These values, attached to different states of health, are obtained from surveys of different populations, where people are polled on the value that they give to the quality of life attributed to symptoms or illness.

First attempts at a QALY

The first attempts to provide a value to life quality were undertaken in 1972, creating the Rosser Index.³ This index based quality of life on observed disability and subjective distress. Disability was split into eight categories (I–VIII) ranging from no disability to unconsciousness. Distress was split into four categories (A–D) ranging from no distress to mild, moderate or severe psychological distress. This

categorization led to the description of 29 patient groups, each with a quality of life value, with the elimination of groups VIII B, C and D. These three groups are not 'useful', as distress cannot be assessed in an unconscious patient. As the Rosser Index was applied, it became apparent that there was significant variation in patients' perceptions of the severity of symptoms. It became clear that patients who were not suffering from a symptom would overestimate its detrimental effect on their quality of life compared to those who were suffering the symptom. The importance of the group making judgments on life quality cannot be minimized. It is evident that there are also substantial differences between the perceptions held by medical professionals and patients. Attempts have been made to avoid observer bias by assessing QALY values in much larger population groups to normalize these discrepancies. It could be argued, of course, that this broadening of the population base may prevent the QALY being effectively applied to the specifics of an individual patient's suffering.

Second attempts at a QALY

The tool most commonly used by NICE for calculating quality of life is the EuroQol EQ-5D questionnaire, however it is not used universally nor is it widely accepted.⁴ The EQ-5D includes assessments of mobility, self-care, daily activities undertaken, pain, discomfort, anxiety and depression. Completion of the EuroQol questionnaire leads to the establishment of 243 distinct states, analogous to the Rosser index groups, each with corresponding numerical quality of life value.⁵ To provide further mystery and enhance confusion, the EQ-5D also includes a visual analogue scale assessment of health; where the subject is asked to rate their health from 0 to 100. Such scales appear to be a simple method of assessing quality of life, but simplicity does not equate to precision, and these scales are among the most subjective ways of assessing life quality.

Other available methods for assessing quality of life include Time Trade Off (TTO) questionnaires. The format of the TTO questionnaire is as follows:

Imagine that you are told that you have 10 years left to live. In connection with this you are also told that you can choose to live these 10 years in your current health state or that you can choose to give up some life years to live for a shorter period in full health. Indicate with a cross on the line the number of years in full health that you think is of equal value to 10 years in your current health state.⁶

The patient is asked to mark the line. If he marks the line at 8 years a result follows which gives a value or score of 0.8 quality of life.

Calculating the QALY

So far we have discussed the difficulties arising from the variety of methods used to assess quality of life, but the reader will be pleased to learn that in practise QALYs can be made even more complex. The basic idea of a QALY is straightforward, with the amount of time spent in a particular health state weighted by the utility score given to that health state. Thus, 1 year spent in 'perfect health' equates to one QALY, while 1 year spent in a state of health valued at 0.25 equates to a quarter of a QALY.

A QALY can then be used to compare one treatment with another, or to evaluate the use of a particular treatment against no intervention. For the purposes of calculating the QALY associated with either of these circumstances, the life expectancy resulting from each scenario, in years, is multiplied by the quality of life factor for each of those years and is expressed as a figure. This data can also be represented graphically, with the quality of life score on the 'y-axis' and time in years on the 'x-axis' to give a value in the form of an area under the QALY-time curve. Both methods can lead to tenebrous Kafkaesque contortions. Let us consider an example of such methodology.

Two years living at 0.6 quality of life results in 1.2 QALY; this is deemed to be the same as living 3 years at 0.4 quality of life. Such calculations will often be a compound result that takes into account a gradual decline in quality of life over time. For instance a treatment might result in 2 years at 0.8 quality of life, 1 year at 0.5 and 1 year at 0.2, the total for the treatment gives a QALY value of 2.3.

Those readers interested in detail will be pleased to learn that the QALY result may be further refined in assessing the specific gain resulting from treatment. To assess the benefit of a treatment the QALY associated with not receiving a treatment is subtracted from that associated with receiving that treatment. This provides an estimate of the equivalent number of years of perfect health gained by the patient as a result of receiving the treatment.

Cost-efficacy

Having used the QALY measurement to compare how much someone's life can be extended and improved by an intervention, NICE then consider cost effectiveness. NICE define this as 'the cost of using a drug to provide a year of the best quality of life

available' and then go on to state 'it could be one person receiving one QALY, but is more likely to be a number of people receiving a proportion of a QALY—for example 20 people receiving 0.05 of a QALY'. This statement is confusing to say the least but also appears to allow a large degree of manipulation of its calculations and, by doing this, NICE covers its processes with a blanket of seeming objectivity.

NICE reviews the QALY gained by a specific treatment and compares this with an existing standard therapy or no intervention. The relative costs of the two interventions are then compared and the cost of the new treatment in excess of relative standard therapy is calculated. This figure is then divided by the QALY gained by the new treatment to give the cost per QALY (£ per QALY). Therefore, a QALY can be used to provide a 'common currency', in the form of a cost utility ratio by which one can compare the cost-effectiveness of one treatment with another, or to evaluate the cost of a particular treatment against no intervention.

The cost per QALY is a critical value that NICE takes into account when deciding whether or not the proposed new treatments can be covered by the NHS. The threshold currently in place is that treatments with a cost per QALY of up to £30 000 are likely to be considered for approval for funding, but those treatments that are more costly are highly unlikely to be authorized for use by the NHS.²

Details, details

Let us consider detail in more detail! What is the reality of the way that specific drugs are dealt with by NICE? The average time for drug development is 14 years, and presto, a new drug becomes available to treat an illness. The availability of new treatments is monitored by the Department of Health and by the National Horizon Scanning Centre. But, the agenda button for review by NICE is pressed by the Department of Health, which sets the schedule for the time course of NICE's decision making process. So, NICE proceeds to work out the QALY, as follows:

Say that patients receiving the drug can expect an average life expectancy of 5 years, at a quality of life of 0.7. With a current standard therapy, life expectancy is just 1 year at a quality of life of 0.3. This is how the new drug's QALY calculation is worked out:

- The QALY of the new drug is $5 \times 0.7 = 3.5$
- Now the QALY of standard therapy is $1 \times 0.3 = 0.3$
- The QALY gained is $3.5 - 0.3 = 3.2$
- The new drug leads to a life expectancy of 5 years and costs £5000, so the cost per patient will be £25 000.

Let us suppose that standard therapy will cost £1000 for the 1 remaining year of life.

- This leads to a calculation that the excess cost of the new drug is £24 000. Therefore, the cost per QALY of the new drug is $£24\,000 \div 3.2$, which works out at £7500 per QALY.
- If this drug were being considered by NICE, it is likely that it would be accepted on the basis of the cost per QALY being $<£30\,000$.

However, one only has to read the 'consideration of evidence section' of a NICE technology appraisal document to realize the process is nowhere near so straightforward. Take the example of the recent NICE appraisal on sorafenib for hepatocellular carcinoma. This was based mainly on data from the Sorafenib Hepatocellular Carcinoma (HCC) Assessment Randomized Protocol study, comparing sorafenib plus best supportive care with placebo versus best supportive care. The ultimate conclusion was that Sorafenib is not recommended for the treatment of advanced HCC in patients for whom surgical or locoregional therapies have failed or are not suitable, but that people currently receiving sorafenib for the treatment of advanced HCC should have the option to continue treatment until they and their clinician consider it appropriate to stop. However the number of different economic models and subgroup analyses used to come to this conclusion is truly mind boggling and one must assume that if different models were used a completely different cost per QALY may be achievable.

The downside

It is clear that a value has to be given to a particular treatment, and that it is impossible to provide everything to everyone, regardless of price. We have finite resources and there is infinite need, but we would strongly argue that there are problems that arise from the use of the QALY in assessing the value of a treatment. While quantity of life expressed in terms of survival, is easy to define and has few problems for comparison, quality of life is much more difficult to evaluate as it encompasses a wide range of different aspects of patients' lives, not just their health status. Quality of life means different things to different people, depending on their life experiences and personal circumstances, and varying with their point in their lives. Patients who suffer a disability unconnected with their medical condition requiring treatment may be disadvantaged. Those who have a disability would be considered to have a lower quality of life and therefore would benefit less from treatment for an independent separate condition than those who, with treatment could be returned to full health.^{7,8}

Furthermore, the effects of a patient's health on the quality of life of others, such as carers or family, do not figure into these calculations but can have wide ranging economic consequences on the state. Additionally, QALYs do not take into account the personal response of individuals to their illness and their views of their need for treatment. The whole concept of a QALY assumes that quality of life is potentially more important than length of life alone. While for many people this concept may hold true, a number of patients are willing to accept many debilitating side effects related to a drug or symptoms from their disease in exchange for a prolongation of life. This is particularly important to those patients wishing to survive in order to be present at or experience a certain event such as the birth of a child or a family wedding or graduation ceremony.

Some people consider QALYs to be ageist because providing treatment for younger people is likely to give a better QALY calculation. Individuals who have a greater life expectancy, typically a younger person, would gain more from treatment on the basis of their predicted life expectancy in comparison with an older person who would have fewer years to benefit. This would not apply to all treatments. If the treatment needs to be continued indefinitely then the cost per QALY calculation may not favour the younger person. The opposing view is that QALYs are not ageist because the age of the patients is not taken directly into account. An emergency life-saving treatment for a 75-year old, who was then expected to live 5 years would have the same priority, under the QALY method, as a 40-year old with a terminal illness whose life-expectancy was 5 years. Others argue that QALYs 'are not ageist enough', as a 40-year old should have greater priority than an 80-year old on the basis that the older person has already had the advantage of a long life and that treatment should be preferentially given to a younger person.⁹ However this concept is extremely dangerous and potentially opens Pandora's Box in terms of just who should be entitled to treatment on the NHS over others.

It is essential that resources are fairly distributed across the populations of the needy and who are we to say, as doctors, that it is more important to treat macular degeneration than diabetes or Alzheimer's rather than breast cancer. However, as we hope we have shown in this article, the QALY is a subjective measure and not an objective tool, calculated on the basis of the cost of a standard treatment, and dependent upon the relative efficacy of the standard treatment and the benefit that the new treatment has over the standard. To ensure fair distribution of

health resources, the cost of standard treatment and the relative benefit of the new treatment over the standard must be taken into account. The prevalence of a disease in a population may also alter how much we are willing to spend on its treatment. For instance, therapies for two different conditions may have the same cost per QALY, however if one condition is very common and the other rare, the total cost associated with funding treatment for the common condition would be much higher. This may discourage organizations such as NICE from authorizing the treatment for the common condition when that for the rare condition may be endorsed. This may result in the cost per QALY threshold being changed for more common conditions, creating healthcare inequality.

Money

A possible flaw in the cost per QALY system is that it has the potential to provide support for less effective treatments. If a new treatment is less effective than standard therapy, but costs much less there will be a cost per QALY benefit in adopting it. This would obviously be a detrimental course of events which we must hope will never be considered.

A potential solution to the problems associated with the QALY lies with value based pricing, which sets selling prices on the perceived value of a product to the customer, rather than on its actual cost to design and manufacture. Historically, Britain's Pharmaceutical Price Regulation Scheme has set maximum and minimum profit levels from the sale of branded drugs to the NHS, allowing companies' freedom to set prices as they please on new substances, but restricting subsequent price increases. This has often lead to inflated drug pricing as companies strive to gain as much profit from a drug while it is still 'in-patent'. The goal of value-based pricing is to align price with value delivered, so the NHS would pay pharmaceutical companies based on the value of a drug to the patient base. This would involve independent cost effectiveness studies that would enable the government or academics to determine a drug's price in conjunction with the pharmaceutical industry. Pricing could be allowed to vary by subgroup since people with certain diseases may benefit more than others. It should lead to a reallocation of revenue from less to more valuable products leading to a reduction in price for some drugs. However, a number of problems exist with this system such as who will perform these analyses as whether it will be impartial or manipulable by interested parties. Drug companies that have invested in the UK may chose to relocate in the fear of potential loss of profits and

the development of more innovative technologies may be stunted due to fear of eventual low profit margins if the drug was deemed relatively invaluable to patients.

Conclusion

The use of QALYs provides a standardized means of comparing the outcome of multiple medical treatments. This is used as a standard tool to ensure that the best possible healthcare is provided for the general public. We have shown that the QALY can be manipulated providing values that depend more on the weather than then any objective parameter. However, many policymakers remain content to apply QALYs uncritically and unreservedly for healthcare technology assessment and despite decades of research we appear no nearer to a solution to a problem that will become increasingly apparent in the age of technological medical advancement. We need to look again at more rational methods of assessing drug efficacy rather than accept that the QALY is a rational standard. This is certain to require the involvement of healthcare professionals, health economic specialists and most importantly patients and carers. Furthermore, a large responsibility must also lie with the pharmaceutical industry to ensure drugs are provided at the right cost in line with the healthcare system to which they provide.

Conflict of interest: None declared.

References

1. Motzer RJ, Hutson TE, Tomczak P, Michaelson MD, Bukowski RM, Rixe O, *et al.* Sunitinib versus interferon alfa in metastatic renal-cell carcinoma. *N Engl J Med* 2007; **356**:115–24.
2. Choosing Health: Making healthy choices easier Government White Paper 16th November 2004.
3. Rosser RM, Watts VC. The measurement of hospital output. *Int J Epidemiol* 1972; **1**:361–8.
4. Stein K, Fry A, Round A, Milne R, Brazier J. What value health?: a review of health state values used in early technology assessments for NICE. *Appl Health Econ Health Policy* 2005; **4**:219–28. Review.
5. Szende A, Oppe M, Devlin N, eds, *EQ-5D Value Sets: Inventory, Comparative Review and User Guide*. UK, Springer, 2007.
6. Burström K, Johannesson M, Diderichsen F. A comparison of individual and social time trade-off values for health states in the general population. *Health Policy* 2006; **76**:359–70.
7. Harris J. QALYfying the value of life. *J Med Ethics* 1987; **13**:117–23.
8. Harris J. Double jeopardy and the veil of ignorance - a reply. *J Med Ethics* 1995; **21**:144–50.
9. Lockwood M. Quality of life and resource allocation. In: Bell JM, Mendus S, eds, *Philosophy and Medical Welfare*. Press Syndicate of the University of Cambridge, 1989:33–5.