

# Ethics and resource allocation: can health care outcomes be QALYfied?

*Femi Oyebode*

**This paper discusses the theoretical foundation of QALY and examines the assumptions which underlie these theories. It argues that there are methodological flaws in the construction of QALY and that there are inherent risks in its possible application to psychiatry. It also draws attention to fundamental ethical problems with the concept of QALY as a tool for valuing the quality of life or well-being of persons.**

Quality adjusted life years (QALY) incorporates both life expectancy and quality of life in the measurement and valuation of the benefits of health care (Williams, 1985). It is a concept which has arisen in the context of growing concern about the finitude of health care resources, and it is meant to provide a rational basis for the allocation of scarce resources. This paper will discuss the theoretical foundation of QALY and then proceed to a critical appraisal of its underlying assumptions and methodology. It will argue that there are a number of methodological flaws in the construction of QALYs, especially in its possible application to psychiatry. Furthermore, it will hope to show that even without these flaws, there are fundamental ethical difficulties with the concept of QALYs as a tool for valuing the quality of life or well being of persons.

## The theory of QALY

The proponents of QALY argue that in a climate of limited resources, decisions need to be made to determine health care priorities, and, that these decisions should at least be based on both the costs of resource inputs and on the health care outcome for the patients involved. The impact of medical intervention upon quality of life must also be a component of any measure of outcome. The QALY measure, it is argued, fulfils such a role (Williams, 1985).

The assessment of quality of life used in QALY is based upon a classification of illness states (Rosser & Watts, 1972). This classification is a general classification of morbidity which is said

to be applicable to all diseases and therefore will permit comparisons of the impact of diseases across disease classes and across medical specialities. It is constructed from a summation of disability and distress questionnaires respectively. There are eight degrees of disability and four degrees of distress which are combined into a valuation matrix, consisting of 32 possible cells, each representing a particular illness state. For example, cell IA represents having no disability and no distress, cell IIB represents slight social disability and mild distress, and cell VC represents inability to undertake any paid employment or inability to continue education or old people confined to home except for outings and short walks and unable to do shopping or housewives able only to perform a few simple tasks and who are moderately distressed.

The weighting for each cell in the valuation matrix was calculated from the responses of 70 individuals comprising ten doctors who were members of a royal college, ten experienced psychiatric nurses, ten experienced general hospital nurses, 20 healthy volunteers, ten patients from medical wards and ten patients from psychiatric wards. The weighting derived from these interviews is a ratio scale in which being healthy is rated 1 and being dead is rated 0. Thus, an illness state valued at 0.5 is believed to be only half the value of being healthy. The assumption here is that two years life expectancy in this state is of equal value to one year of healthy life.

In clinical situations, the technique may provide the following kind of information: intervention A which produces perfect health for ten years yields a QALY score of 10, whereas intervention B which extends life for 20 years but at a quality valued at 0.2 yields a QALY score of 4. Cost per QALY can also be derived: if intervention A costs £20 000 then cost per QALY is 2000, whereas if intervention B costs £1000, then cost per QALY is 250. In this scenario priority must be given to intervention B.

### Methodological problems

The QALY is derived from the responses of 70 individuals. There is no reason to believe that these individuals are representative of the general public or of doctors, nurses or indeed of patients. The valuation arrived at from these interviews is dependent upon the values of the respondents. Therefore, in order to establish a robust and consistent valuation, the method of sample selection must be explicit and valid. There is empirical evidence that when the sample was enlarged to 300 respondents, the valuation of certain illness states altered significantly (Spiegelhalter *et al.*, 1992). This suggests that the method of sample selection and the composition of the sample may affect the valuation of illness states arrived at.

The method of calculating the QALY values each future year according to a fixed discount rate of 5%. This method is drawn from economics where future costs are discounted to their net present value to reflect the fact that delays in expenditure reduce the real cost. This is a way of saying that goods which cost £5 today will be worth much less in five years, if there is a constant inflationary rate of 5% in the given period. This is of course only strictly true where there is rising inflation. Within the QALY context, the assumption is that years in the distant future are not worth as much as in the immediate future (Spiegelhalter *et al.*, 1992). This assumption is open to question. Time preference may differ significantly between different age cohorts as indeed it may between individuals. Young people may place a lower weight on good health now than on good health a few years ahead when raising a family may be the paramount desire. The weight placed on good health in years beyond that when children have grown up may then fall. There is little information about the strength of preference of individuals in different age groups and therefore no good reason to believe that a fixed discount rate is appropriate or correct. The level of discount used also alters the resultant QALY and the level chosen does not appear to be based upon explicit grounds.

QALY assumes that there is a constant proportional trade-off between length of life and health status. This entails an individual being prepared to sacrifice some proportion of his or her remaining years of life in order to achieve a given improvement in his or her health status, irrespective of the absolute number of years that remain, for example 12 years in excellent health against 15 years in poor health. There is some evidence that people are only willing to make this kind of trade-off where the absolute length of time spent in poor health is greater than five years (McNeil *et al.*, 1981).

QALY also assumes risk neutrality especially when its proponents state that a beneficial health care activity is one that generates a larger, positive amount of QALYs. However, it is conceivable that a majority of a population might not only prefer but also regard a prospect which offers each of them as individuals a larger probability of a smaller QALY gain to the prospect of a smaller probability of a larger QALY gain as beneficial, and they might be willing to pay a kind of risk premium by choosing the prospect which in QALY terms offers a lower expected value. The converse is also true. There may be a demand for some treatments like liver transplantation which do poorly by the cost per QALY criterion but which offer dramatic benefits for at least some of those treated.

Some illness states were valued by the respondents as worse than death. However, people who actually experience these states may not regard their lives as valueless or as having a negative value. This fact was made clear in the use of QALY in the Oregon experiment where respondents who had experienced some illness states rated them as of higher quality than people who had not. Yet the views of these disabled people were excluded from the QALY calculations. This suggests that "if the people whose quality of life is ranked low do not themselves share that judgement of their condition, then something must be wrong with the claim that what is being ranked is the real quality of people's lives" (Menzel, 1992).

There are problems with the application of QALYs to psychiatry. The emphasis on the extension of life as one of the major components of outcome is probably not appropriate to psychiatry (Wilkinson *et al.*, 1990; Oyeboode *et al.*, 1992). As a method, QALYs, is likely to favour treatment of acute fatal conditions which effect a full recovery characterised by a good quality of life. Treatments for fatal or non-fatal conditions which produce only minimal or no extension of life and moderate improvement in quality of life may be unjustifiably disfavoured. Treatments for conditions such as schizophrenia may fall into this category, although treatments for the major mood disorders which can be shown to prevent fatality (suicide) may be favoured (Boyle & Callahan, 1993). Nonetheless, there may still be a formalised structural disadvantage for psychiatric treatments if QALY were to be accepted as a method for allocating resources.

### Ethical considerations

The preceding discussion can very easily give the impression that if only a more elegant mathematical or statistical technique existed, the use of QALY would cease to be problematic. This paper will now argue that there are fundamental

problems of an ethical nature with the use of QALY and indeed, that some of the issues which were examined above cannot be solved in a technical way precisely because they are located in a moral space. In other words, that there are questions which relate to our values, our understanding of the good and the good life. In addition, our conception of how to distribute the benefits and burdens of social cooperation are also involved.

Life expectancy is often taken to be a measure of well-being. This is because we assume that a prospect of ten years is roughly twice as good as the prospect of five years. But, this is far from clear. One's life projects, what one accomplishes in life are clearly central to the value of life. Health state and the length of life are not the sole and immutable determinants of what gives value and quality to life (Loomes & McKenzie, 1989).

What we want to know is whether well-being is a homogenous value which is easily represented on a single dimension and comparable both within and between individuals. There are differing desires and needs, the satisfaction of which may result in a sense of well-being. It is possible, in some situations, to say that a certain amount of one thing is more satisfying than another amount of another thing. However, it is also conceivable that no increase in one kind of value can overtake another kind of value. In other words, that the two values cannot be compared quantitatively at all, that they are incommensurable. The question is "are interpersonal comparisons of well-being possible?" The quality of life of a cripple cannot easily be compared with that of an individual who has ten years to live with leukaemia or with that of a person with schizophrenia who has 30 years to live. We lack knowledge about not just how much individuals get out of some good, but also what their conception of the good is. However, some authorities like Griffin (1986) argue that the idea that values are incomparable is not strictly true. For example, he claims that when we despair of ranking artists of diverse achievements, perhaps two novelists, we may be able to say that one is dry but full of insight and the other is unperceptive but hilarious. This, in his view, is a sort of comparability. In my view, this is the wrong example to use. It would be difficult if not impossible to decide the relative merits of Picasso and Eliot. And, this would be because the achievements are within different domains and therefore incommensurable. The essence of this argument is that "there is a plurality of values, that they clash, that they all matter . . . that there are no permanent orderings or rankings among them . . . and that persons may go in very different directions and still lead equally valuable lives" (Griffin, 1986).

QALY is located in the utilitarian tradition where it is assumed that an individual ought to choose between alternatives so as to maximise utility. In this case, this depends on survival duration, health status, and cost of intervention. In addition, QALY is derived from an aggregation of individual values which are assumed to represent societal values and interests. This position is fundamental to classical utilitarianism where 'everybody counts for one, nobody for more than one'. This is the principle of equal respect but the problem is that one person's well-being may be sacrificed merely to increase the aggregate social utility. As Rawls (1971) puts it "utilitarianism does not take seriously the distinction between persons" and "it mistakes impersonality for impartiality". Harris (1987) develops this argument further when he draws attention to the fact "it does not follow that where the choice is between three years of discomfort for me or immediate death on the one hand, and one year of health for you, or immediate death on the other, that I am somehow committed to the judgement that you ought to be saved rather than me". He argues that such a policy does not value life or lives at all.

### Comment

QALY is likely to be discriminatory against the elderly and the mentally and physically disabled and it is likely to perpetuate structural disadvantages associated with gender and race (Harris, 1987), while maximising benefits for the already advantaged. The danger of QALY to discriminate against disabled people in Oregon was recognised by the US government (Menzel, 1992).

The proponents of QALY argue that 'its use is justified by the lack of alternative tools of analysis'. The assumption is 'better to use a flawed technique that gives visibility to the criteria being used than to muddle through'. This is a mistaken view. QALYs in health care decision making has as much potential for welfare gain or loss as any new pharmaceutical compound and should be stringently evaluated (Klein, 1989). In conclusion, there is no compelling reason to believe that the QALY in its present form is not fundamentally flawed. It may, with modification, serve a useful but minor role in the resource allocation process.

### References

- BOYLE, P.J. & CALLAGHAN, D. (1993) Minds and hearts: priorities in mental health services. *Hastings Center Report*, 23, no. 5, Special Supplement.
- GRIFFIN, J. (1986) *Well-Being: its meaning, measurement, and moral importance*. Oxford: Oxford University Press.

- KLEIN, R. (1989) The role of health economics. *British Medical Journal*, **229**, 275–276.
- HARRIS, J. (1987) QALYfying the value of life. *Journal of Medical Ethics*, **13**, 117–123.
- LOOMES, G. & MCKENZIE, L. (1989) The use of QALYs in health care decision making. *Social Science & Medicine*, **238**, 299–308.
- MCNEIL, B.J., WEICHELBAUM, R. & PAUKER, S.G. (1981) Speech and survival: tradeoffs between quality and quantity of life in laryngeal cancer. *New England Journal of Medicine*, **305**, 982–987.
- MENZEL, P.T. (1992) Oregon's denial: disabilities and quality of life. *Hastings Center Report*, **22**, 21–25.
- OYEBODE, F., CUMELLA, S., GARDEN, G. & NICHOLLS, J. (1992) Development of outcome measures in acute psychiatry. *Psychiatric Bulletin*, **16**, 618–619.
- RAWLS, J. (1971) *A Theory of Justice*. Oxford: Oxford University Press.
- ROSSER, R. & WATTS, V.C. (1972) The measurement of hospital output. *International Journal of Epidemiology*, **1**, 361–368.
- SPIEGELHALTER, D.J., GORE, S.M., FITZPATRICK, R., FLETCHER, A.E. *et al* (1992) Quality of life measures in health care. III: resource allocation. *British Medical Journal*, **305**, 1205–1209.
- WILKINSON, G., CROFT-JEFFREYS, C., KREKORIAN, H. *et al* (1990) QALYs in psychiatric care. *Psychiatric Bulletin*, **14**, 582–585.
- WILLIAMS, A. (1985) The value of QALYs. *Health & Social Services Journal*, **XCIV**, 4957.

Femi Oyeboode, Consultant Psychiatrist, Queen Elizabeth Psychiatric Hospital, Mindelsohn Way, Edgbaston, Birmingham B15 2QZ

# The pattern of delays in Mental Health Review Tribunals

Stephen Blumenthal and Simon Wessely

**Mental Health Review Tribunals (MHRTs) are intended to protect the liberty of the detained patient. In practice, however, they have been said to fall short of providing this safeguard. This is frequently for administrative reasons. We report a systematic examination of the pattern of delays in holding MHRTs. Our main finding is that delay cannot be attributed to a single factor but rather reflects the inherent complexity of some cases.**

It is estimated that it takes an average of 24 weeks to arrange Tribunals for applicants who are detained under a restriction order and 15 weeks for non-restricted patients (Department of Health, 1993). For a patient detained on a six month order (section 3 of the Mental Health Act, 1983) Tribunals are often held at the end of their period of detention when the patient may be close to release anyway, the delay therefore negating the purpose of the Tribunal.

## The study

The study was divided into two phases. The first, retrospective study, involved an examination of 150 Tribunal applications selected at random from the four regional Tribunal offices which had taken place over the past five years. The second, prospective stage, was a more detailed examin-

ation of factors identified in the first phase. Two hundred Tribunal applications were randomly selected from the four regional Tribunal offices. They were tagged upon receipt and tracked until completion which meant withdrawal of application, discharge prior to the hearing, or the communication of a decision. Section 2 applications were excluded as there are no problems with delay in these cases. In both phases of the study we intentionally oversampled special hospital cases in order that half the sample were taken from the NHS and half from special hospitals.

Social workers and responsible medical officers involved in each of the 200 tagged prospective Tribunals were identified and sent a questionnaire and we examined the effect of responses on the time taken to the submission of reports. Information was also collected from each of the patients' files at the MHRT offices.

We used survival analysis to determine the extent to which various factors contributed to Tribunal delay. The relevant statistic is a hazard ratio, both crude and adjusted. The relative size of the hazard ratio is a measure of the extent to which the particular variable contributes to delay.