How to do (or not to do) ...  

Calculating QALYs, comparing QALY and DALY calculations

FRANCO SASSI
Department of Social Policy, The London School of Economics and Political Science, London, UK

Quality-adjusted life years (QALYs) have been used in the assessment of health interventions for three decades. The popularity of the QALY approach has been constantly increasing, although the debate on its theoretical underpinnings and practical implications is still ongoing. Disability-adjusted life years (DALYs), also widely debated, were shaped some 20 years later, broadly within the same conceptual framework but with a number of important differences.

This paper provides a comprehensive formulation of QALY calculation methods, offering practical instruments for assessing the impact of health interventions, similar to those made available elsewhere for calculating DALYs. Systematic differences between QALYs and DALYs are explained by reference to two examples: the prevention of tuberculosis and the treatment of bipolar depression. When a health intervention is aimed at preventing or treating a non-fatal disease, the relationship between QALYs gained and DALYs saved depends on age of onset and duration of the disease, as well as the quality of life and disability weights. In the case of a potentially fatal disease, a larger number of factors may determine differences between outcomes assessed with the two metrics. The relative importance of some of these factors is discussed and illustrated graphically in the paper. Understanding similarities and differences between QALYs and DALYs is important to researchers and policy makers, for a sound interpretation of the evidence on the outcomes of health interventions.

Key words: quality-adjusted life years, disability-adjusted life years, outcome assessment, cost-effectiveness analysis

Introduction

The term ‘quality-adjusted life year’ (QALY) was first used in 1976 by Zeckhauser and Shepard to indicate a health outcome measurement unit that combines duration and quality of life (Zeckhauser and Shepard 1976). But the underlying concept had been formally shaped in the early 1970s in the development of a ‘health status index’ (Fanshel and Bush 1970; Torrance 1970; Torrance et al. 1972), while an earlier study of the treatment of chronic renal disease (Klarman et al. 1968) had used a subjective adjustment for quality of life. Early applications of the health status index include one on tuberculin screening (Bush et al. 1972) and one on screening for phenylketonuria (Bush et al. 1973). The underlying assumptions of the QALY model were spelled out by Pliskin et al. (1980), who demonstrated that the QALY maximization criterion is justified in a multi-attribute utility theory framework under the following conditions: utility independence between life years and health status; constant proportional trade-off; and risk neutrality on life years. These conditions and the utility theory foundations of QALYs were further discussed in a number of contributions, including those of Myamoto and Eraker (1985), Loomes and McKenzie (1989), Mehrez and Gafni (1989). An extensive review published in 1992 counted 51 economic evaluations using QALYs as the outcome measure (Gerard 1992). Only a few years later the QALY framework was widely accepted as the reference standard in cost-effectiveness analysis (Gold et al. 1996; McPake et al. 2002; Drummond et al. 2005), amid a continuing debate on its theoretical underpinnings and practical implications (e.g. Bleichrodt and Johannesson 1996). Today, QALYs are used in most economic evaluations, and by many regulatory agencies which have made cost-effectiveness analysis an integral part of their decision-making processes.

The QALY framework provided a basis for the development of a number of health outcome measures, including the disability-adjusted life year (DALY) in the early 1990s. The DALY is primarily a measure of disease burden (disability weights measure loss of functioning) but its use in cost-effectiveness analysis is also relatively common, and this paper is concerned with the latter. As a measure of outcome in economic evaluation, the DALY differs from the QALY in a number of aspects. Most importantly, the DALY incorporates an age-weighting...
function assigning different weights to life years lived at different ages, and the origins of disability and quality of life weights differ significantly. Although the disability profiles upon which DALY calculations are based tend to be simple (e.g. a constant disability is often assumed), the actual calculations may be relatively complicated, as illustrated in some detail by Murray (1994) and, in this Journal, by Fox-Rushby and Hanson (2001). On the other hand, quality of life profiles (or health profiles) for QALY calculations tend to be more elaborate, allowing for sequential upward or downward health status changes over time, but the corresponding calculation methods can be made less cumbersome by using a discrete approximation of a continuous health function (Drummond et al. 2005). Perhaps the only comprehensive mathematical formulation of a QALY-type measure, the ‘health status unit years’, was produced by Torrance (1976) with reference to a population perspective.

This paper illustrates the methods for calculating QALYs, providing formulas that can be applied directly by researchers, similar to those made available elsewhere for DALY calculations. The paper also includes two practical examples, in which the results of QALY and DALY calculations are compared and systematic differences are shown. This paper is about calculation methods, and it does not aim at providing a comprehensive discussion of the conceptual and methodological differences between the two measures, which are well documented in other contributions. In particular, Broome (1993) provides a detailed discussion of the conceptual framework of QALYs, while methods for eliciting health state utility values are presented in Torrance (1986). Readers may refer to Loomes and McKenzie (1989) and Mehrrez and Gafni (1989) for a critique of the utility framework on which QALYs are based, and to Sassi et al. (2001) for a review of the ethical and distributional issues involved. Most of the challenges to the QALY framework have been based on the difficulties involved in making interpersonal comparisons and aggregating individual utilities; the assumptions on which health utility elicitation methods are based; and the implicit discrimination against the elderly and the chronically ill or disabled. The debate on the conceptual framework, key assumptions and ethical implications of DALYs is illustrated in Anand and Hanson (1997), Murray and Acharya (1997), Williams (1999) and Arnesen and Nord (1999). Key challenges to the DALY framework have focused on the equity implications of age-weighting and of the standard life expectancy assumption used in cross-country comparisons, but also on the methods used to assess disability weights. A direct comparison of the two measures is presented in Gold et al. (2002).

Calculating QALYs

QALYs are used primarily to correct someone’s life expectancy based on the levels of health-related quality of life they are predicted to experience throughout the course of their life, or part of it. The number of QALYs lived by an individual in one year is simply:

\[ Q \text{ALYs lived in one year} = 1 \times Q \text{ with } Q \leq 1; \]

where \( Q \) is the health-related quality of life weight attached to the relevant year of life. From this descends that someone’s quality-adjusted life expectancy (QALE) at age \( a \) can be defined as:

\[ \text{QALE} = \sum_{i=a}^{a+L} Q_i \]

where \( L \) is the residual life expectancy of the individual at age \( a \), and \( i \) represents individual years within that life expectancy range. If someone's quality of life is predicted to change over shorter than yearly periods, \( i \) can be taken to represent correspondingly shorter units of time, such as a month, a week or even a day. In these cases, \( L \) will have to be defined consistently. When time preference, and thus discounting, is incorporated into the equation, QALE becomes:

\[ \text{Discounted QALE} = \sum_{i=a}^{a+L} \frac{Q_i}{(1 + r)^{t-a}} \]

where \( r \) is the discount rate.

However, QALYs are rarely used to simply assess someone’s quality-adjusted life expectancy. The main use of QALYs is within the framework of cost-effectiveness analysis, to assess the improvement in quality-adjusted life expectancy obtained through a specific health intervention (i) relative to a situation in which either no intervention or a standard alternative intervention is provided. In such analysis, the number of QALYs gained can be determined as follows:

\[ \text{QALYs gained} = \sum_{i=a}^{a+L} \frac{Q_i}{(1 + r)^{t-a}} - \sum_{i=a}^{a+L} \frac{Q_i}{(1 + r)^{t-a}} \]

where \( Q_i \) is a vector of health-related quality of life weights predicted (or observed) for each time period \( t \) following the intervention. When QALY calculations are undertaken for the purpose of assessing the QALY gain following an intervention, the focus is on the time period during which an individual is affected by a disease, or by the effects of its treatment. Therefore \( L \) should be defined as the duration of the disease, while \( L' \) is the period over which the individual enjoys the benefits of treatment (or possibly suffers the adverse consequences of it). Normally, the period \( L' \) will be at least as long as \( L \), but it will be longer than \( L \) when treatment extends the individual’s life expectancy (in this case, \( L' \) will correspond to the individual’s entire life expectancy with treatment), or when treatment may negatively affect the individual’s quality of life for a period longer than \( L \) (in this case, \( L' \) will correspond to the entire period over which treatment affects the individual’s quality of life).

Health-related quality of life weights used in QALY calculations differ from disability weights (\( D \)) used in
Daly calculations in several respects. Although measured on similar scales, the former represent levels of quality of life enjoyed by individuals in particular health states, while the latter represent levels of loss of functioning caused by diseases. The former are normally measured on a scale in which 1 represents full health and 0 represents death, therefore higher values correspond to more desirable states and states deemed worse than death can take negative values. The latter are measured on a scale in which 0 represents no disability, therefore lower scores correspond to more desirable states. The two types of weights are also derived in different ways, using different elicitation techniques and different groups of subjects.

In practice, DALY calculations tend to be based on a universal set of standard weights based on expert valuations, while QALY calculations often rely on preference-based health-related quality of life measures directly elicited from general population samples or from groups of patients. The most common preference elicitation techniques are the standard gamble and the time trade-off, both choice-based (Torrance 1986). These may be applied directly, or indirectly in the assessment of the value of individual dimensions of multi-attribute systems like the Health Utilities Index (Torrance et al. 1996) or the EuroQol (Dolan 1997).

QALYs do not incorporate an age-weighting function. Therefore, one QALY has always the same value, regardless of the age at which it is lived, although this does not imply neutrality over age distributions (Sassi et al. 2001). Discounting procedures are also different in QALYs and DALYs (discrete the first, continuous the second). If the DALY discounting procedure is applied, the QALE formula can be revised as:

\[
QALE = \int_{x=a}^{a+L} Q e^{-r(x-a)} dx = Q \frac{1 - e^{-rL}}{r} \tag{I}
\]

where \( e \) is Napier’s mathematical constant and \( x \) is the individual’s age. The corresponding formula for QALYs gained follows from this:

\[
\text{QALYs gained} = Q^i \frac{1 - e^{-rL^i}}{r} - Q \frac{1 - e^{-rL}}{r} \tag{II}
\]

where \( L^i \) and \( Q^i \) are, respectively, the period over which treatment affects the individual’s quality of life, and the quality of life weight with treatment; while \( L \) and \( Q \) are the corresponding parameters without treatment. This situation is illustrated in Figure 1, in which the QALY gain is the area between the health profiles.

When treatment does not affect the duration of the disease, but only the individual’s quality of life, i.e. when \( L^i = L \), equation (II) becomes:

\[
\text{QALYs gained} = (Q^i - Q) \frac{1 - e^{-rL}}{r} \tag{III}
\]

\[\text{Figure 1. Health profiles with constant quality of life. Note: Health profiles with intervention } i \text{ (solid line), and without intervention (broken line).}\]

Equations (I) to (III) are based on the assumption that health-related quality of life remains constant throughout the individual’s residual life expectancy, or disease duration. While this assumption is common in DALY calculations, it is much less so in QALY calculations, in which health-related quality of life is normally allowed to vary with disease progression. A more general formula for calculating QALE, which allows for quality of life to vary over time, can be developed by assuming that the individual’s residual life expectancy (\( L \)) is divided into \( N \) consecutive time periods \( n_m \) (with \( 1 \leq m \leq N \)), each upper-delimited by time point \( t_m \), whereby \( t_N = a + L \), and each characterized by a level of health-related quality of life \( Q_m \). The time periods \( n_m \) may be of different durations. Based on this information, the QALE formula can be re-written as:

\[
\text{QALE} = \sum_{m=1}^{N} Q_m \frac{e^{-r(t_m-a)} - e^{-r(t_{m-1}-a)}}{r} \quad \text{with } t_0 = a \tag{IV}
\]

The formula for calculating the number of QALYs gained through an intervention \( i \) follows directly from the above:

\[
\text{QALYs gained} = \sum_{p=1}^{P} Q_p \frac{e^{-r(t_{p}-a)} - e^{-r(t_{p-1}-a)}}{r} - \sum_{m=1}^{N} Q_m \frac{e^{-r(t_m-a)} - e^{-r(t_{m-1}-a)}}{r} \tag{V}
\]

where the life expectancy with the intervention (\( L^i \)) is divided into \( P \) time periods \( n_p \), defined in the same way as the \( n_m \) above, and \( Q^i_p \) is a vector of health-related quality of life weights predicted (or observed) for each time period \( n_p \) following the intervention, as illustrated in Figure 2.

**Comparing QALYs and DALYs: practical examples**

The calculation methods illustrated in the previous section will be applied in two examples, one on tuberculosis,
Therefore:

The number of QALYs an individual will live while affected by the disease can be determined using equation (III) as follows:

\[
\text{QALYs gained} = (1 - 0.726) \times \frac{1 - e^{-0.03 \times 0.5}}{0.03} = 0.14
\]

The corresponding number of GBD DALYs saved, with disease onset at age 45, would have been 0.17. The QALY gain would be greater if the expected duration of the disease were longer than 6 months. Figure 3 shows the number of QALYs gained by preventing one case of tuberculosis, as a function of the expected duration of the disease had it not been prevented. The same figure also shows what the corresponding numbers of DALYs saved would be, depending on the age of onset of the disease, had this not been prevented. For convenience and ease of comparison it has been assumed that \( D = 0.726 \) for all ages. The figure shows that, for most age groups, numbers of QALYs gained and DALYs saved tend to diverge progressively as disease duration becomes longer. However, the ratios between QALYs gained and DALYs saved tend to be relatively stable across different disease durations, and are insensitive to the value of \( Q \) or \( D \) (as long as \( Q = 1 - D \)). This allows the calculation of conversion factors indicating the extent of the divergence between the two measures (as illustrated in Table 1), which are valid only under the restrictive assumptions previously discussed. Conversion factors vary by age of disease onset and by disease duration. Discount rate variations have a very limited impact on them. Under the assumptions described, the number of DALYs saved is equal to the number of QALYs gained multiplied by the relevant conversion factor \((C_{45,0.5} = 1.228)\).

The impact of relaxing the assumption \( Q = 1 - D \) can be assessed by using appropriate quality of life weights for tuberculosis. Dion et al. (2002) report a mean standard gamble value of 0.68 for moderate disease. If this was used
in QALY calculations, the QALY gain would change to 0.16, and this would vary with disease duration as illustrated by the relevant curve in Figure 3.

**A potentially fatal condition**

Our second example refers to a chronic disease affecting both quality and duration of life, and is based on a case described by Fox-Rushby and Hanson (2001) to illustrate DALY calculations. A Chilean woman becomes affected by bipolar depression at age 35. In the absence of treatment, this woman would live a further 10 years with a disability (D) of 0.6 and then die. The woman’s quality-adjusted life expectancy (QALE) at the age of disease onset can be calculated using equation (I). It is initially assumed that \( Q = 1 - D = 0.4 \). The life expectancy lost is simply ignored, as QALYs focus on the duration and quality of life of the years actually lived by individuals.

\[
\text{QALE} = 0.4 \times \frac{1 - e^{-0.03 \times 10}}{0.03} = 3.46
\]

If treatment were available, the woman would be able to live her entire residual life expectancy, with a disability reduced to 0.302 for the rest of her life. Female life expectancy in Chile at the age of 35 is 44.13 years, therefore:

\[
\text{QALE with treatment} = 0.698 \times \frac{1 - e^{-0.03 \times 44.13}}{0.03} = 17.08
\]

The QALY gain is the difference between the woman’s QALE with and without treatment – as in (II):

\[
\text{QALYs gained} = 17.08 - 3.46 = 13.62
\]

This is only marginally different from the number of DALYs saved, which in this example is 13.72. Given the assumptions made here, and given the use of the same discounting procedure as in DALY calculations, the difference is entirely attributable to the age weighting factor (included in DALY calculations but not in QALY calculations).

When \( L' \neq L \), the relationship between QALYs gained and DALYs saved is determined by a large number of parameters, including: the quality of life, or disability, weights with and without treatment; the age of onset of the disease; the duration of the disease with and without treatment. Therefore, it is not possible to calculate conversion factors like those reported in Table 1.

If the assumption that \( Q = 1 - D \) is relaxed, appropriate quality of life weights can be used to assess the value of \( Q \). Tsevat et al. (2000) report a mean (standard gamble) utility of 0.77 in a sample of patients undergoing various types of treatment. Applying this in QALY calculations leads to a QALY gain figure of 15.38 (instead of 13.62) in the baseline case. Figures 4–6 illustrate how QALYs gained and DALYs saved vary in relation to changes in, respectively, age of disease onset (\( a \)), duration of disability without treatment (\( L \)), and disability weight with treatment (\( D' \)). These show that when \( Q \) is equal or very close to 1 – D, the factor potentially generating the largest divergence between the two measures is age at disease onset, while other factors have limited or no impact. However, even relatively small departures of the value of \( Q \) from 1 – D may determine substantial differences between QALYs gained and DALYs saved.

**Discussion**

This paper provides an illustration of calculation methods for assessing quality-adjusted life expectancy and for measuring the outcomes of health interventions in terms of QALYs. Two examples in different disease areas have
shown that age of disease onset is an important factor determining variations between numbers of QALYs gained and DALYs saved, when interventions are compared using the two metrics. The pattern of variation is mostly dictated by the shape of the age-weighting function. QALYs gained exceed DALYs saved when disease starts in the very early years of life and is of short duration; when the disease starts in later years, up to young adulthood, DALYs saved exceed QALYs gained, sometimes by a relative large margin; finally, when the disease starts in late adulthood and in older ages, QALYs gained again exceed DALYs saved. These conclusions are based on the use of the age-weighting function originally proposed in the GBD study (Murray and Lopez 1996), still most widely applied in DALY calculations. Results would have been different if based on a different function, or if QALYs had been age-weighted too, as advocated by some (see Sassi et al. 2001 for a discussion of the latter).

The examples have also shown that differences between quality of life and disability weights may cause further divergence between QALYs gained and DALYs saved. In some cases, estimates of the loss of quality of life used in QALY calculations may be very close, or equal, to disability estimates used in DALY calculations. However, variations can often be expected in either direction. In our examples, we have used quality of life weights derived from the literature to illustrate the possible extent of such variations.

The examples in this paper are based on the assumption that the assessment of the relevant interventions is country-specific. Instead, the original formulation of DALYs for the GBD study was aimed at supporting cost-effectiveness comparisons on a global scale, therefore a standard life expectancy was assumed in order not to disadvantage populations with a shorter actual life expectancy. The two approaches may lead to different results, an example being an intervention that avoids premature mortality caused by a given disease (as in the second example above). The standard life expectancy assumption leads to a consistently larger estimate of DALYs saved, and the difference is greater where actual life expectancy is shorter.

Although QALYs and DALYs stem from the same broad conceptual framework, they are not interchangeable, as they are partly based on different assumptions and different methodologies (for instance, methods for eliciting quality of life and disability scores). Understanding systematic differences between the two measures is important for enabling policy makers to form a sound judgement on the existing evidence about the outcomes of health interventions.
Endnote

1 Fox-Rushby and Hanson indicate the slightly different figure of 13.81 DALYs saved. This is because, in their calculation of Years of Life Lost (YLL), Fox-Rushby and Hanson determine the loss of life expectancy (L) as the ‘standard expectation of life at age [of death]’, rather than the expectation of life at the time of disease onset minus the number of years lived with disability.

Acknowledgements

The author wishes to thank Mrigesh Bhatia for discussions that led to an earlier version of this paper. The usual disclaimer applies.

References


Biography

Franco Sassi is a senior lecturer in Health Policy in the Department of Social Policy at the London School of Economics and Political Science (LSE). He received his PhD from the London School of Hygiene and Tropical Medicine (LSHTM). His research interests focus on the economic analysis of health services, including the evaluation of diagnostic and screening interventions, multi-attribute utility assessment techniques for measuring health outcomes, measures of equity in economic analysis, and inequalities in health and access to health care.

Correspondence: Franco Sassi, PhD, Department of Social Policy, The London School of Economics and Political Science, Houghton Street, London, WC2A 2AE, UK. E-mail: f.sassi@lse.ac.uk