

## Editorial

Is the EQ-5D fit for purpose in mental health?<sup>†</sup>

John Brazier

**Summary**

The EQ-5D is a widely used questionnaire for calculating quality-adjusted life-years (QALYs) for assessing cost-effectiveness in healthcare. It reflects the impact of common mental health conditions such as mild to moderate depression but seems to be more problematic for use in

people with psychotic and severe and complex non-psychotic disorders.

**Declaration of interest**

None.

John Brazier has more than 20 years' experience of conducting economic evaluations for policy makers and of methods of valuing health for economic evaluation.

The National Institute for Health and Clinical Excellence (NICE) in England and others agencies around the world use cost-effectiveness to inform resource allocation decisions in healthcare.<sup>1</sup> Interventions are assessed in terms of their cost per quality-adjusted life-year (QALY). The QALY is designed to permit comparisons across programmes of care, including mental health. In this issue of the *Journal*, Saarni and colleagues<sup>2</sup> claim their evidence shows that one of the main instruments used to calculate QALYs, the EQ-5D, is problematic for use in psychotic disorders.

**EQ-5D**

The number of QALYs is calculated by multiplying each time period by the health-related quality of life associated with that period on a scale of zero (for dead) to one (full health) (states worse than dead such as vegetative states may be given a negative value). In the context of a clinical trial with multiple follow-ups, the number of QALYs for each person is calculated as the area under the curve, with a horizontal axis for time measured in years and a vertical axis indicating their 'health state value'. Although there have been objections to the QALY, including theoretical and philosophical, it provides a way of measuring the benefits of different interventions on a common simple metric.<sup>3</sup>

Quality-adjusted life-years require a value for health on a scale of 0–1 and one instrument for doing this is the EQ-5D.<sup>4</sup> This patient-reported outcome measure has five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression, and each dimension has three levels (no problem, some problem, severe problem). Together, these five dimensions define a total of 243 health states scored using values obtained from a survey of the general population. Over 3000 members of the public were asked how many years of their life they would be willing to sacrifice to avoid any given ill health state and live in good health. The EQ-5D is a 'preference-based' measure of health and although there are other preference-based measures using different dimensions and different methods of valuation, it is the most widely used in healthcare.

It is claimed that the EQ-5D is applicable to all interventions and patient groups. This claim has support for many physical

conditions where these instruments have managed to pass psychometric tests of reliability and validity. For other conditions, such as in relation to visual impairment in macular degeneration and to hearing loss,<sup>4</sup> the claim has not been substantiated.

**EQ-5D in mental health**

In mental health, evidence is rather limited but suggests a potentially mixed picture. There is evidence that generic instruments are able to reflect the impact of common conditions such as mild to moderate depression and anxiety,<sup>5</sup> but in a study of chronic schizophrenia using measures of psychopathology and functioning to establish change, the EQ-5D did not have a significant correlation with negative symptoms, disorganisation, depression, excitement and general symptoms.<sup>6</sup> The impact of a range of mental disorders on scores on the generic preference-based SF-6D has been modelled using data from 8580 respondents from the Office for National Statistics Psychiatric Morbidity Survey.<sup>7</sup> After adjusting for covariates, major anxiety disorders and depressive episodes were found to have a significant impact on SF-6D scores. However, obsessive compulsive disorder, personality disorder and probable psychosis were not significant.

Saarni and colleagues present the results of a screening survey of psychosis in a large representative sample of the Finnish population using the EQ-5D alongside another generic preference-based measure not widely used outside of Finland, the 15D.<sup>8</sup> Schizophrenia, other non-affective disorders and affective psychotic disorders were all associated with lower EQ-5D scores compared with scores for the non-psychotic population. However, in contrast to the other measures, the EQ-5D index did not show a statistically significant reduction for participants with delusional or bipolar I disorders. Another interesting finding was that there were no statistically significant reductions in the EQ-5D after controlling for depression, although schizophrenia and schizoaffective disorders were associated with significant reductions in the 15D. In general, quality-of-life measures did not correlate well with symptoms or clinician-assessed outcomes except in the case of depression. The authors conclude that this poses a challenge for economic evaluation since interventions typically target positive symptom reduction that would be missed by measures such as the EQ-5D.

A recent study by Barton *et al* in 77 participants with psychosis found differences in EQ-5D values between groups with mild and more severe illness and improvements post-intervention.<sup>9</sup> They interpreted this as supporting the use of EQ-5D in this condition. However, all significant differences were between groups defined

<sup>†</sup>See pp. 386–394, this issue.

by measures of depression and only one functioning scale. There were no significant differences in EQ-5D values between groups defined by positive and negative syndrome, general quality of life and social and occupational functioning assessment. These findings are quite similar to those of Saarni and colleagues.

Such evidence does not prove that the EQ-5D is invalid in these populations. Indeed, it is not possible to obtain definitive proof for a concept such as self-reported health-related quality of life since there is no gold standard. Tests of validity have to assume that individuals with psychosis should have lower scores owing to sociodemographic factors and other symptoms such as depression. Weak correlations with clinical assessments also do not provide conclusive evidence. However, it would seem highly likely that psychotic disorders will induce feelings of fear and stigma that have far-reaching consequences that are not captured by depression alone. Measures need to be tested and in the health measurement field, researchers have long recognised that the best we can do is to examine concepts such as content validity (e.g. the extent to which the content reflects the impact of mental health problems on quality of life) and construct validity (e.g. the extent to which the scores reflect known differences between groups).

The content of the EQ-5D was developed from literature reviews and expert judgement. Although this approach may be efficient in terms of time and useful where a consensus is required, the problem is that it might not reflect what matters to patients. The alternative approach would be to generate items from patients. Across the mental health conditions, this approach is all the more important because the outward signs and symptoms of the condition may poorly reflect the impact on the patient's quality of life from their point of view. The need to involve patients in the development and testing of measures has even been recognised by the US Food and Drug Administration in its guidance on the development of patient-reported outcome measures.<sup>10</sup> A study funded by the Medical Research Council currently being undertaken in Sheffield aims to fill this gap by examining the content validity of the EQ-5D (and SF-36) by undertaking in-depth interviews with mental health service users.

There are many outcome measures specific to mental health that make an important contribution to clinical research.<sup>11</sup> However, these condition-specific measures are not suitable for use in economic evaluation since they are not preference-based, that is they have not been scored with the values of the general public obtained using a recognised elicitation technique (as required by NICE and similar agencies around the world).

### Where do we go from here?

It is important to further test generic measures such as the EQ-5D in mental health. There are considerable advantages to using measures that assess physical and mental health problems together. Where generic measures are found to be adequate, as may be the case in some common mental health problems, NICE has recommended estimating ways to map from condition-specific measures (e.g. PHQ-9) onto EQ-5D in order to predict EQ-5D scores from the instrument used in any given trial or study. However, this depends on the generic instrument being appropriate or else it will miss some specific dimensions being picked up by the condition-specific measure.

It is unlikely that generic measures will be adequate for all mental health conditions. Therefore it will be necessary to start

a programme of work to develop mental health-specific preference-based measures. In some areas of mental health there has already been an impressive amount of work to develop quality-of-life instruments such as the PHQ-9 and CORE-OM. Given that such measures exist and have been widely tested, it may be possible to build a preference-based measure using them by applying modern psychometric techniques to help simplify them. This approach has been used recently with the CORE-OM.<sup>12</sup>

For some conditions (e.g. psychotic and personality disorders), there may be a case for developing a new preference-based measure that reflects the views of mental health services users and at the same time passes standard psychometric testing. It should be developed from in-depth qualitative interviews and psychometrically tested. The second stage would be to value health states defined by the new descriptive system using one of a number of possible techniques, such as time trade-off,<sup>4</sup> using the values of the general public (to satisfy NICE) and mental health service users. This would enable health economics to better meet the challenge posed by Saarni and colleagues.

**John Brazier**, PhD, MSc, BA, Health Economics and Decision Science, School of Health and Related Research, University of Sheffield, Regent Court, 30 Regent Street, Sheffield S10 2FD, UK. Email: j.e.brazier@sheffield.ac.uk

First received 13 May 2010, final revision 30 Jul 2010, accepted 19 Aug 2010

## References

- 1 National Institute for Health and Clinical Excellence. *Guide to the Methods of Technology Appraisal*. NICE, 2008 (<http://www.nice.org.uk/media/B52/A7/TAMethodsGuideUpdatedJune2008.pdf>).
- 2 Saarni SI, Viertiö S, Perälä J, Koskinen S, Lönnqvist J, Suvisaari J. Quality of life of people with schizophrenia, bipolar disorder and other psychotic disorders. *Br J Psychiatry* 2010; **197**: 386–94.
- 3 Brazier JE, Ratcliffe J, Tsuchiya A, Solomon J. *Measuring and Valuing Health for Economic Evaluation*. Oxford University Press, 2007.
- 4 Brooks RG. Euroqol: the current state of play. *Health Policy* 1996; **37**: 53–72.
- 5 Lamers LM, Bouwmans CA, van Straten A, Donker MC, Hakkaart L. Comparison of EQ-5D and SF-6D utilities in mental health. *Health Econ* 2006; **15**: 1229–36.
- 6 van de Willige G, Wiersma D, Nienhuis FJ, Jenner JA. Changes in quality of life in chronic psychiatric patients: a comparison between EQ-5D and WHOQoL. *Qual Life Res* 2005; **14**: 441–51.
- 7 Brazier J. Measuring and valuing mental health for economic evaluation. *J Health Serv Res Policy* 2008; **13** (suppl 3): 70–5.
- 9 Barton GR, Hodgkins J, Mugford M, Jones PB, Croudace T, Fowler D. Measuring the benefits of treatment for psychosis: validity and responsiveness of the EQ-5D. *Br J Psychiatry* 2009; **195**: 170–7.
- 10 U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research, U.S. Department of Health and Human Services FDA Center for Biologics Evaluation and Research, U.S. Department of Health and Human Services FDA Center for Devices and Radiological Health. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health Qual Life Outcomes* 2006; **4**: 79.
- 11 Gilbody S, House A, Sheldon T. Routine administration of health related quality of life and needs assessment instruments to improve psychological outcome. *Psychol Med* 2002; **32**: 1345–56.
- 12 Mavranzouli I, Brazier JE, Young TA, Barkham M. *Using Rasch analysis to form plausible health states amenable to valuation: the development of CORE-6D from CORE-OM in order to elicit preferences for common mental health problems*. Health Economics and Decision Science Discussion Paper 09/09. University of Sheffield, 2009 (<http://www.shf.ac.uk/content/1/c6/01/87/47/HEDS%20DP%200909.pdf>).