Article:
Brazier, J.E. and Fitzpatrick, R. (2002) Measures of Health Related quality of life in an imperfect world: a comment on Dowie. Health Economics, 11 (1). pp. 17-19. ISSN 1057-9230

https://doi.org/10.1002/hec.669

HEALTH ECONOMICS
Health Econ. **11**: 17–19 (2002)
DOI:10.1002/hec.669

DEBATE

Measures of health-related quality of life in an imperfect world: a comment on Dowie

John Brazier*[,1] and Ray Fitzpatrick[2]
[1] *Sheffield Health Economics Group, School of Health and Related Research, The University of Sheffield, UK*
[2] *Department of Public Health, Institute of Health Sciences, University of Oxford, UK*

Professor Dowie has written an interesting and thought provoking paper on a long lasting debate in the literature on measuring health related quality of life. The debate between generic and condition specific measures (CSMs) has not progressed a great deal with time and he is right to question a purely psychometric approach that currently tends to focuses on effect sizes. He has also presented an interesting challenge to the compromise solution suggested by a number of psychometricians to adopt both types of measure.

The paper is predicated on the apparently useful distinction between knowledge- versus decision-validity, but this is something of a straw man argument. Advocates of CMSs are not interested in knowledge-validity but are concerned with different types of decisions. Furthermore, producers of criteria for judging the merits of measures (including ourselves [1,2]) do not presume it to be an absolute question of whether a measure is 'valid' but hope such lists assist researchers and other users to find the best measure for their purpose, such as informing clinical or resource allocation decisions.

For patients and clinicians important information is provided by a trial through the use of CSMs in terms of the impact on the problem that they bring to the doctor – i.e. proximal concerns. Neither the patient nor the doctor are necessarily anticipating an impact on the more distal dimensions typically covered by generic measures. Professor Dowie's response to this narrow view is that there may be important consequences of the treatment that are not covered by the CSM and consequences for co-morbidities. This is the basis of the usual argument for using both types of measure in trials.

Professor Dowie goes on to argue that a generic measure 'is intended to cover the full range of health outcomes' whereas the condition specific is by definition intended to cover a narrower range. On the basis of this distinction of intention he argues that in most cases the GEN should be used alone, but he accepts that there may be circumstances where a CSM is preferable, but that it is never preferable to use both (for decision-making purposes). We would just like to make a few comments about this distinction of intention.

CSMs typically focus on fewer domain of interest and hence may be able to ask more questions about each domain and thereby achieve a greater degree of refinement. For example, the EQ-5D has a large jump in scores between perfect health and any level of illness of 1.0–0.88 and has only two levels of imperfect health for each dimension. However, Professor Dowie seems sceptical as to the empirical evidence regarding the sensitivity of CSMs over generics. We agree there are cases that do not support this common claim in the psychometric literature [3], but equally there are many where this has been found including: EQ-5D in chronic pulmonary disease [4] and in cosmetic surgery [5], the SF-36 in urinary incontinence [6] and the example of urinary incontinence that he provides [7].

*Correspondence to: Sheffield Health Economics Group, School of Health and Related Research, The University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA, UK. Tel.: 0114 222 0715; e-mail: j.e.brazier@sheffield.ac.uk

The main concern people have with generic measures is that their domains are too far removed from those of the patient and are often irrelevant to the condition. Even at the level of intention it is not clear to us that developers of generic measures intend to cover the full range of health problems. Developers themselves have used terms such as 'core' [8] or described the content of instruments as addressing the 'majority' of health concerns of people [9]. A better approach would be to conceive of a matrix of health, where one specific generic measure covers Sections A and B, one specific CSM measure covers B and C and so on. The EQ-5D, for example, does not directly assess cognition, seeing, hearing, sleep, energy or breathlessness, but these are often the primary problems of concern to patients

The choice between generic measures and CSMs should be made primarily on the appropriateness of the content of the instrument. Given the increasing emphasis on patient's views then an obvious group to turn to for advice are those whom we expect to benefit (or suffer) from the intervention, rather than impose our own views. A researcher will also want to choose the measure (or measures) that best covers the domains within the resource available for the study (including the time of patients or their proxies). This requires some empirical consideration of the responsiveness of the measure. Professor Dowie is right to criticise the psychometric literature on responsiveness, that is often driven by comparisons of effect sizes. Biggest is not always best. What is required is to examine the ability of measures to respond to minimally important changes as reported by patients in health transition questions [10]. Sensitivity is important since many trials are examining small and hard to find differences between active treatments, particularly in the pharmaceuticals field.

For social decisions, it might be thought that the case for a reliance on generic preference-based measures is much stronger. In order to include health related quality of life in an economic evaluation, the only rigorous way would be to conduct a cost utility analysis (CUA) which requires a preference-based measure and since only (some) generic measure are preference-based then this would seem to rule out all CSMs.

At a social level, the choice of appropriate domains is more than simply a matter for patients. Although their views will be important, there is an additional concern with what represents an appropriate benefit from public funds. At least in the UK, the reluctance to fully meet patient demands for viagra seems to be partly based on this. Similarly debates over IVF services have been concerned partly concerned with the appropriateness of funding such services from health care budgets. However, we would contend that most of the health domains covered by CSMs and generics would be regarded as within the legitimate domain of public funding provided there was a health related cause.

One or other type of measure may prove to be unnecessary since it has not picked up any changes or because the generic has been found to adequately reflect the changes in the CSM. There will be scenarios, however, where it will be necessary to combine the two measures. Professor Dowie is concerned with the resultant intuitive rather analytical basis of any decision, since this can only be done within the framework of a cost-consequences analysis and this is not usually regarded as a technique of economic evaluation at all [11]. We would be more concerned by a policy of deliberately excluding domains of health important to patients from consideration of economic appraisal on the grounds of analytical inconvenience.

We agree that in the longer term the solution will be to either improve the generic measures by making them more sensitive and in some case more relevant by adding dimensions. It should also be recognised that generics vary in the domains they cover and a more appropriate choice of generic may solve the problem in some cases. The HUI-III [12], for example, covers the dimensions of cognition, seeing and hearing excluded by the EQ-5D and SF-6D but not usual activities dimension found in the EQ-5D, and neither have a dimension for energy, such as the SF-6D [13]. Another solution in some cases may be to improve CSMs and extend their coverage to treatment outcomes. For social decision-making it is also possible to estimate preference-weighted CSMs (for example see the work of Revicki and colleagues [14], although this raises technical problems of combining generic and CSM preference-based measures where the preference function is non-additive). For the short term, clinical researchers and health economists must work within the domain of a considerably less than perfect world and this may sometimes mean using both types of measure.

References

1. Fitzpatrick R, Davey C, Buxton M, Jones D. Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assessment* 1998; **2**(14): 1–74.
2. Brazier J, Deverill M, Green C, Harper R, Booth A. A review of the use of health status measures in economic evaluation. *Health Technol Assessment* 1999; **3**(9): 1–164.
3. Fitzpatrick R, Ziebland S, Jenkinson C, Mowat A. A comparison of the sensitivity to change of several health status instruments in rheumatoid arthritis. *J Rheumatol* 1993; **20**: 429–436.
4. Harper R, Brazier JE, Waterhouse JC, Walters SJ, Jones NMB, Howard P. A comparison of outcome measures for patients with chronic obstructive pulmonary disease (COPD) in an outpatient setting. *Thorax* 1997; **52**: 879–887.
5. Klassen A, Fitzpatrick R, Jenkinson C, Goodacre T. Contrasting evidence of the effectiveness of cosmetic surgery from two health related quality of life measures. *J Epidemiol Community Health* 1999; **53**: 440–441.
6. Kobelt et al. Quality of Life aspects of the over active bladder and the effect of treatment with tolterodine. *Brit J Urol* 1999; **83**: 583–590.
7. Jenkinson C, Gray A, Doll H, Lawrence K, Keoghane S, Layte R. Evaluation of index and profile measures of health status in a randomised controlled trial: comparison of the MOS SF-36, EuroQol, and disease specific measures. *Med Care* 1997; 1109–1118.
8. Ware JE, Snow KK, Kolinski M, Gandeck B. 1993. *SF-36 Health Survey Manual and Interpretation Guide*. The Health Institute, New England Medical Centre, Boston, MA.
9. Williams, A. (1995). The measurement and valuation of health: a chronicle. Centre for Health Economics. *Discussion Paper*, University of York.
10. MacKenzie R, Charlson M, DiGiola D, Kelley K. A patient-specific measure of change in maximal function. *Arch Intern Med* 1986; **146**: 1325–29.
11. Drummond MF, Stoddart GL, Torrance GW. *Methods for the Economic Evaluation of Health Care Programmes*. Oxford Medical Publications: Oxford, 1997.
12. Torrance GW, Furlong W, Feeny D, Boyle M. Multi-attribute preference functions. Health Utilities Index. *Pharmaco Econ* 1995; **7**, 503–520.
13. Brazier JE, Harper R, Thomas K, Jones N, Underwood T. Deriving a preference based single index measure from the SF-36 *J. Clin Epidemiol* 1998; **51**(11): 1115–1129.
14. Revicki DA, Leidy NK, Brennan-Diemer F, Sorenson S, Togias A. Intergrating patients preferences into health outcomes assessment: the multiattribute asthma symptom utility index. *Chest* 1998; **114**(4): 998–1007.

