Comparing Cost-Utility Estimates
Does the Choice of EQ-5D or SF-6D Matter?

Tracey H. Sach, PhD,†‡ Garry R. Barton, PhD,† Claire Jenkinson, BSc,‡ Michael Doherty, MD,§ Anthony J. Avery, MD,‡ and Kenneth R. Muir, PhD‡*

Background: A number of different measures can be used within cost-utility analyses, we compared results according to both the EQ-5D and SF-6D.

Methods: A randomized trial was conducted to compare 4 options for people with knee pain. Over the 2 year trial period, the change in cost to health-service was estimated, and both the EQ-5D and SF-6D were used to estimate the change in quality-adjusted life years (QALYs). Using a complete case analysis, the cost-utility (incremental cost-effectiveness ratio [ICER]) of each option, according to both the EQ-5D and SF-6D, was calculated and assessed in relation to the cost-effectiveness threshold of £20,000 per QALY.

Results: Of the 389 participants, 247 had complete cost, EQ-5D and SF-6D data. According to the EQ-5D, option 1 had an estimated ICER of £10,815 (compared with option 4), option 2 was dominated by option 1, and option 3 was subject to extended dominance. Conversely, according to the SF-6D, option 3 had an ICER of £9999 (compared with option 4), option 2 had an ICER of £36,883 (compared with option 3), and option 1 was subject to extended dominance.

Conclusion: The EQ-5D and SF-6D estimated that different options (1 and 3, respectively) were cost-effective at the £20,000 per QALY threshold, demonstrating that the choice of measure does matter.

Key Words: cost-utility, cost-effectiveness, EQ-5D, SF-6D

(*Med Care* 2009;47: 889–894)

As resources are scarce economists have developed a common utility scale (where 0 equates to death and 1 to full health) on which to compare the benefits of many interventions.[1] Most commonly, utility scores are estimated using different health state descriptive measures such as the EQ-5D,[2] health utilities index,[3] or SF-6D.[4] However, as these measures are based on both different health state descriptions,[5] and valuation techniques,[6] they have been shown to produce different utility scores for the same group of patients.[7–11] Though a number of studies have compared the utility scores produced by different measures, few studies have assessed the impact that the choice of measure has on estimates of cost-utility, and it has been argued that further research is required in this area.[12,13] This article aims to fulfil this remit by comparing cost-utility estimates for both the EQ-5D and SF-6D using data from a randomized trial which compared 4 options for overweight and obese people who had knee pain.

From the *School of Chemical Sciences and Pharmacy, University of East Anglia, Norwich, United Kingdom; †Health Economics Group, School of Medicine, Health Policy and Practice, University of East Anglia, Norwich, United Kingdom; ‡School of Community Health Sciences, University of Nottingham, Nottingham, United Kingdom; and §Academic Rheumatology, University of Nottingham, Nottingham, United Kingdom.

The LIKP study was supported by the UK Arthritis Research Campaign (ARC) grant (13550).

Reprints: Garry Barton, PhD, Lecturer in Health Economics, Health Economics Group, School of Medicine, Health Policy and Practice, University of East Anglia, Norwich, NR4 7TJ, United Kingdom. E-mail: g.barton@uea.ac.uk.

ISSN: 0025-7079/09/4708-0889

METHODS

Participants

All participants were taking part in the Lifestyle Interventions for Knee Pain (LIKP) study—a randomized controlled trial which was designed to compare the effectiveness and cost-effectiveness of 4 different options (1) diet and strengthening exercise advice, (2) dietary advice, (3) strengthening exercise advice, and (4) leaflet provision (hereafter these options are referred to as 1, 2, 3, and 4 as the main focus of this article is methodological). All registered patients in 5 UK general practices were recruited if they were aged ≥45 years, reported knee pain on most days of the last month, had a body mass index >28.0 kg/m^2, and gave consent to be randomized to one of the aforementioned 4 options.

Estimating Costs

The methods used to estimate the change in cost to the health service for each participant in the LIKP study have been described elsewhere,[14] briefly they were as follows. Three cost components were monitored—visits by the health care professionals providing each option, provision of exercise bands to participants in either of the strengthening exercise groups (options 1 and 3), and the change in the cost of analgesics. All costs were estimated over the 2 year trial period at 2006/7 price levels, costs incurred in the second year were discounted a rate of 3.5%,[15] and multiple imputation was used where data were missing.[16]

Estimating Utility

Participants were asked to complete both the EQ-5D and SF-6D at baseline, 6, 12, and 24 months postrandomization. With the EQ-5D the respondent is asked to report the level of problems they have (no problems, some/moderate problems, and severe/extreme problems) with regard to mobility, self-care, usual activities, pain/discomfort, and anxiety/depression.[2] Up to 243 different health state descriptions can be derived from responses to these 5 dimension, and utility scores were assigned to each of these states using the York A1 tariff.[17] SF-6D[4] scores were derived from responses to eleven questions on the SF-36[18] questionnaire. The SF-6D has 6 dimensions (physical functioning, role limitations, social functioning, pain, mental health and vitality) which have between 4 and 6 severity levels. Scores for the 18,000 potential health states were estimated using the consistent version of the SF-6D algorithm.[5]

Estimating Quality-Adjusted Life Years

Using the area under the curve method (with adjustment for baseline differences)[19] the EQ-5D and SF-6D scores at baseline, 6, 12, and 24 months postrandomization were used to estimate the QALY gain (or loss), which accrued over the 2 year trial period for each measure (QALY scores in the second year were discounted at 3.5%).

Analyses

To compare the EQ-5D and SF-6D results, a complete case analysis approach[20] was used. Thereby, only those participants who had an estimated QALY gain according to both measures were included in all subsequent analyses. Summary statistics were calculated, where these included the mean change in cost and the mean QALY gain for each of the 4 options.

Cost-Utility

For each of the 4 options estimates of cost-utility were calculated separately for both the EQ-5D and SF-6D (from the perspective of the health service). Dominated options (those with both a higher mean change in cost and lower mean QALY gain, compared with another option) were excluded, as were options which were subject to extended dominance (as combinations of other options could provide a greater benefit at equivalent cost).[21] Subsequently, the cost-utility of nonexcluded options was calculated by estimating the incremental cost per QALY gain (incremental cost-effectiveness ratio [ICER]) associated with each option, relative to the next best alternative. Finally, in line with guidance by the UK National Institutes of Health and Clinical Excellence,[22] we sought to identify the most cost-effective option at the threshold (λ) of £20,000 per QALY.

Decision Uncertainty

To depict the level of uncertainty associated with the decision as to which option was cost-effective, according to each of the EQ-5D and SF-6D, we constructed the cost-effectiveness acceptability curve (CEAC)[23,24] and the expected value of perfect information (EVPI).[25,26] The CEAC depicts the probability that an option is cost-effective at different levels of λ (ie, according to how much one is willing to pay for a QALY gain). It was constructed using the technique of nonparametric bootstrapping,[27] where 10,000 simulations of the (per participant) cost and effect were drawn for each option (with replacement) from the original cost and effect data. The probability of being cost-effective was then equivalent to the proportion of the 10,000 simulations for which each option had the highest net benefit[28] (net monetary benefit $= \lambda \times E–C$, where E is the QALY gain and C is the cost) at different values of λ.

When making decisions about resource allocation, in addition to estimating the cost-effectiveness of particular options, it has been argued that one should consider whether further research should be undertaken to reduce the level of uncertainty associated with that decision.[29] The EVPI addresses this by estimating the upper bound of the value of undertaking further research.[26] As outlined previously,[23] we estimated it at a λ of £20,000 per QALY by first calculating the value of perfect information in each of the aforementioned simulations (the difference between the highest net benefit across each of the 4 options and the net benefit of the most cost-effective option). Second, the EVPI (per patient) was calculated by taking the expectation of the value of perfect information across each simulation.

RESULTS

Participants

After sending out 12,500 questionnaires, and conducting a local media campaign, a total of 389 participants were both eligible and consented to take part in the LIKP study.

Analyses

The EQ-5D was completed at baseline, 6, 12, and 24 months postrandomization by 270 participants, 254 completed the SF-6D at each of these time points, and 247 completed both measures. Of these, 67 received option 1, compared with 78 for option 2, 47 for option 3, and 55 for option 4. Summary statistics for each option are provided in Table 1. The estimated mean change in cost over the 2 year trial period (for each option), ranged between −£19.38 and £821.93 (the negative value can be explained by the fact that those who received option 4 (leaflet provision) did not receive any related health care visits, but had lower mean analgesic costs at 24 months postrandomization, compared with baseline).

For each option the mean EQ-5D score was higher postrandomization, compared with baseline, where the mean QALY gain over the 2 year trial period ranged between 0.076 (option 3) and 0.150 (option 1) (mentioned in Table 1). The mean SF-6D scores also improved, but to a lesser extent as the mean QALY gain ranged between 0.021 (option 4) and 0.065 (option 2) (mentioned in Table 1). The rank ordering of options (from most to least effective) was however different according to the EQ-5D[1,2,4,3] and SF-6D.[2,1,3,4]

Cost-Utility

By comparing the mean change in cost and mean QALY gain according to the EQ-5D across each option (mentioned in Table 1) it was apparent that option 2 was dominated by option 1, similarly option 3 was dominated by option 4. Compared with option 4, option 1 had a mean incremental cost of £794.90 and a mean incremental QALY gain of 0.074, giving an ICER of £10,814.74. According to

TABLE 1. Estimates of the Mean QALY Gain, Mean Change in Cost and the Incremental Cost Effectiveness Ratio According to the EQ-5D and SF-6D

	Option							
	1		2		3		4	
	EQ-5D	SF-6D	EQ-5D	SF-6D	EQ-5D	SF-6D	EQ-5D	SF-6D
Baseline utility score	0.535	0.649	0.544	0.643	0.600	0.648	0.591	0.678
6 mo utility score	0.657	0.699	0.626	0.697	0.665	0.661	0.623	0.686
12 mo utility score	0.599	0.670	0.608	0.680	0.611	0.675	0.657	0.694
24 mo utility score	0.620	0.674	0.602	0.667	0.672	0.697	0.617	0.688
Mean QALY gain	0.150	0.052	0.116	0.065	0.076	0.050	0.076	0.021
Mean change in cost	£776		£822		£268		-£19	
Incremental cost per QALY (ICER)	£10,815	ED	D	£26,883	D	£9999		

D indicates dominated; ED, subject to extended dominance.

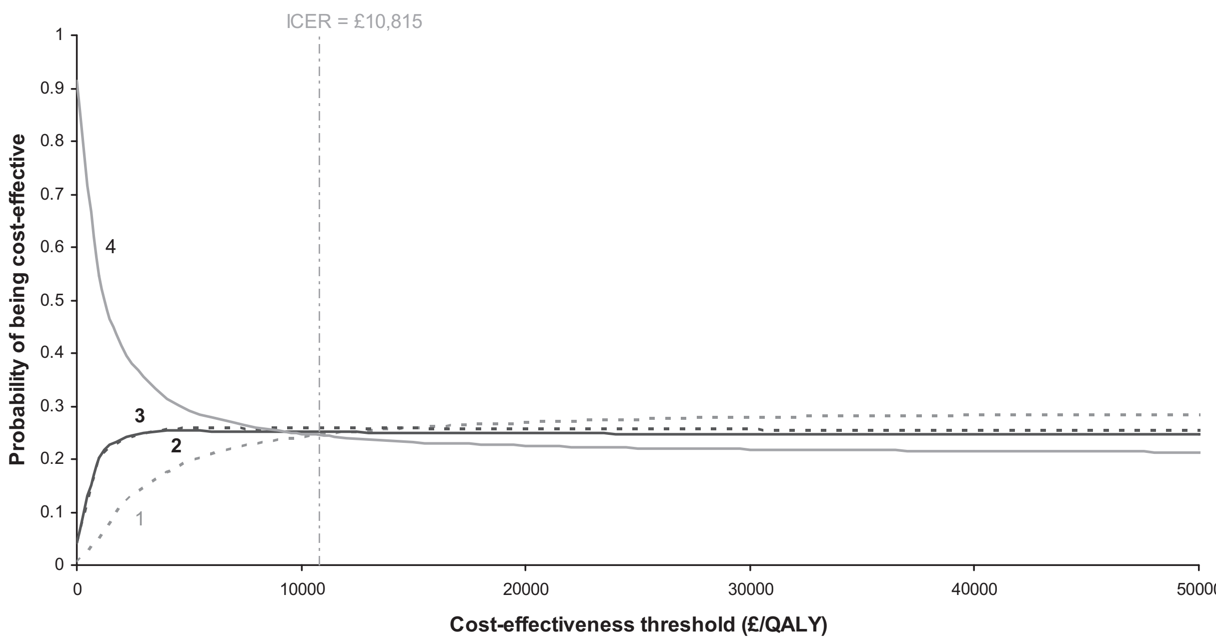

FIGURE 1. Cost-effectiveness acceptability curves (CEACs) for options 1 (light dashed line), 2 (bold continuous line), 3 (bold dashed line), and 4 (light continuous line) constructed using the EQ-5D. The incremental cost-effectiveness ratio (ICER) for option 1 is denoted by the vertical dashed line, for λ values below this ICER option 4 was estimated to be cost-effective, compared with option 1 for higher values of λ.

the SF-6D option 1 was subject to extended dominance as combinations of options 2 and 3 could provide a higher benefit at lower cost, option 3 was estimated to have an ICER of £9999.39, (compared with option 4), and option 2 was estimated to have an ICER of £36,882.58 (compared with option 3) (mentioned in Table 1). Thus, at a λ of £20,000 per QALY option 1 was estimated to be cost-effective according to the EQ-5D, compared with option 3 according to the SF-6D.

Decision Uncertainty

The CEACs which were based on the EQ-5D are plotted in Figure 1. It can be seen that at low levels of λ (<£5000 per QALY) option 4 was estimated to have the highest probability of being cost-effective. However, for values >£5000 per QALY the probability of each option being cost-effective was <30%. This demonstrates that, according to the EQ-5D, there is much uncertainty associated with the decision as to which option is most cost-effective. The CEACs derived from on the SF-6D are plotted in Figure 2, and portray a similar story. With regard to the EVPI at a λ of £20,000 per QALY it was estimated to be £7195 per patient according to the EQ-5D, and £3219 according to the SF-6D (this result is explained further in the Discussion).

DISCUSSION

We have shown that, when comparing the cost-effectiveness of 4 different options for people with knee pain, the EQ-5D and SF-6D estimated that different options (1 and 3, respectively) were most cost-effective.

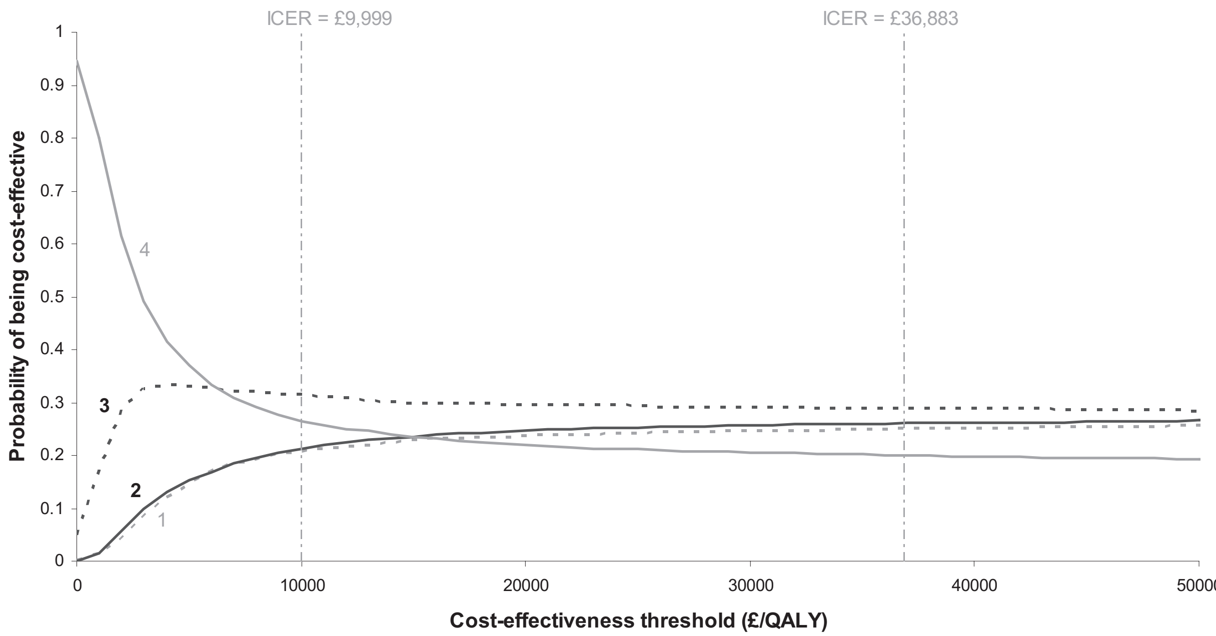

FIGURE 2. Cost-effectiveness acceptability curves (CEACs) for options 1 (light dashed line), 2 (bold continuous line), 3 (bold dashed line), and 4 (light continuous line) constructed using the SF-6D. The incremental cost-effectiveness ratio (ICER) for options 3 and 2, respectively, are denoted by the vertical dashed lines, for λ values below the first ICER option 4 was estimated to be cost-effective, compared with option 3 before the second ICER, and option 2 thereafter.

Utility scores on the SF-6D have a narrower range (0.296 to 1.00),[5] compared with the EQ-5D (−0.594 to 1.00).[17] Additionally, time trade-off valuation scores (the technique used with the EQ-5D[30]) tend to be higher for milder states, but lower for more severe states, compared with standard gamble scores (which are used with the SF-6D).[31] As a result, it has been proposed[6,32] that estimated utility gains will tend to be higher according to the EQ-5D, compared with the SF-6D, as has been shown previously.[7–11,33–35] Our results in Table 1 seem to support this conclusion as the QALY gain associated with each of the 4 options was estimated to be higher according to the EQ-5D. However, the rank ordering of the 4 options (from the least to highest QALY gain) differed between the EQ-5D and SF-6D and as such the ratio of the QALY gain for the EQ-5D relative to that for the SF-6D also differed across options 1 (2.9), 2 (1.8), 3 (1.5), and 4 (3.6). Thus, in spite of the fact that the QALY gain for a particular option was always estimated to be higher for the EQ-5D, the most cost-effective option was estimated to differ according to each of the 2 measures.

The main implication of the finding that estimates of cost-effectiveness differ between measures is that further research is needed to determine which measure is the most appropriate in this, and other, clinical areas. Thereby, building upon previous studies, which have assessed how the different health state descriptions of different measures may account for differing utility scores,[8,10,13,34] future studies should be encouraged to include more than 1 measure, and to test the various criteria that have been developed to assess whether instruments are measuring what they purport to for eg, construct validity and responsiveness.[36,37] Due to the range of issues that need to be considered when choosing between measures[4,36] there is, however, insufficient space in this article to determine which of the EQ-5D and SF-6D is most appropriate for people with knee pain. Indeed, though the EQ-5D could be argued to be the more sensitive measure in this population group (utility gains were consistently estimated to be higher according to the EQ-5D), this does not necessarily mean that it is the preferred measure as this could represent an overestimation of the true gain in utility associated with the particular change in health. Recently, in the light of these factors, and to ensure a consistent approach is taken in different analyses, the UK National Institute of Health and Clinical Excellence has developed a reference case, wherein certain methods for conducting cost-effectiveness analyses are recommended, including use of the EQ-5D to estimate adult utility scores.[22]

An additional finding was that the EVPI (the potential value of further research) was more than twice as high according to the EQ-5D, compared with the SF-6D. The EVPI is determined by both the probability, and consequences, of making the wrong decision.[26] As the probability that each option was cost-effective (depicted by the CEAC) was similar according to the EQ-5D and SF-6D (mentioned in Figs. 1 and 2, respectively), the different EVPI values result from different estimates of the consequences of making the wrong decision. The latter is dependent upon the net benefit ($\lambda \times$ E–C) of each option.[23] Therefore, as the difference in QALYs (between options) was estimated to be higher according to the EQ-5D, the difference in net benefit (between options) was also estimated be higher, as was the consequences of making the wrong decision and the EVPI. As has been argued previously,[23] this demonstrates that both the CEAC and EVPI provide different information about the estimated level of the uncertainty associated with a decision regarding cost-effectiveness, and that both values can be of use to decision makers. With the CEAC uncertainty is characterized by estimating the probability that an option is cost-effective at different levels of the cost-effectiveness threshold, whereas the EVPI estimates the upper limit for the value of undertaking further research.

We are aware of 4 articles which have compared estimates of cost-utility according to different measures.[38–41] Pickard et al[38] estimated the ICER for both a group of asthma and stroke patients, where these were calculated using 10 different algorithms which have previously enabled scores on both the SF-12[42] and SF-36[18] to be converted into utility scores. They found that the ICER for the asthma patients ranged between \$30,769 and \$63,492, compared with \$27,972 and \$72,727 for the stroke patients. Moreover, in line with our results, they found that those algorithms which produced more favorable ICERs for the asthma patients did not necessarily produce more favorable results for the stroke patients.[38] Marra et al[41] compared 2 drug options for patients with rheumatoid arthritis where, as previous studies had not included a utility instrument, mapping[43] was used to convert scores on a clinical measure into estimated utility scores. The ICERs for the EQ-5D (\$46,322), SF-6D (\$69,826), HUI2 (\$53,429), and HUI3 (\$32,018) covered a range >\$37,000 per QALY. van den Hout et al[40] compared 2 therapy programs for people with rheumatoid arthritis. The exercise program was found to be dominated according to the EQ-5D, yet had an ICER of €67,000 per QALY according to the SF-6D. Michaels et al[39] compared conservative treatment to surgery for varicose veins, where the ICER for surgery was estimated to be £4417 according to the EQ-5D and £7175 according to the SF-6D. With the possible exception of the last study, it can be seen that the findings of these studies are in line with ours in that different measures tend to produce different ICERs, which can thereby lead to quite different conclusions about whether an intervention is cost-effective or not.

One potential limitation of this article is that a complete case analysis approach was used. This enabled one to compare the EQ-5D and SF-6D scores, and cost-utility estimates, on the same group of patients without introducing other factors, which might account for any differences. However, as these 247 participants may be unrepresentative of the 389 who took part in the LIKP study, by way of sensitivity analysis, we also used multiple imputations to estimate the cost-effectiveness of the different options. Though the actual results were slightly different (eg, according to the EQ-5D, option 1 was estimated to have an ICER of £10,469 when multiple imputation was used,[14] compared with £10,815 with a complete case analysis [as mentioned in Table 1]), the implications were the same. The main strength of the article is that it is thought to be the first to compare (1) estimates of cost-utility, based on different measures, when evaluating more than 2 options and (2) how the level of uncertainty, according to the CEAC and EVPI, also differs across measures.

CONCLUSION

We have estimated the cost-utility of 4 different options according to both the EQ-5D and SF-6D. The estimated utility scores and QALY gain for each of the 4 options differed between measures, as did the rank ordering of options (from least to most effective). Consequently, we have shown that the choice of measure can greatly influence decisions as to both which option is most cost-effective, and the value of undertaking further research.

REFERENCES

1. Drummond MF, Sculpher MJ, Torrance GW, et al. *Methods for the Economic Evaluation of Health Care Programmes.* 3rd ed. New York, NY: Oxford University Press; 2005.
2. Brooks R. EuroQol: the current state of play. *Health Policy*. 1996;37:53–72.
3. Feeny D, Furlong W, Torrance GW, et al. Multi-attribute and single attribute utility functions for the Health Utilities Index Mark 3 system. *Med Care*. 2002;40:113–128.
4. Brazier JE, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ*. 2002;21:271–292.
5. Brazier JE, Roberts J, Tsuchiya A, et al. A comparison of the EQ-5D and SF-6D across seven patient groups. *Health Econ*. 2004;13:873–884.
6. Tsuchiya A, Brazier J, Roberts J. Comparison of valuation methods used to generate the EQ-5D and the SF-6D value sets. *J Health Econ*. 2006;25:334–346.
7. Conner-Spady B, Suarez-Almazor ME. Variation in the estimation of quality-adjusted life-years by different preference-based instruments. *Med Care*. 2003;41:791–801.
8. Longworth L, Bryan S. An empirical comparison of EQ-5D and SF-6D in liver transplant patients. *Health Econ*. 2003;12:1061–1077.
9. Pickard AS, Johnson JA, Feeny DH. Responsiveness of generic health-related quality of life measures in stroke. *Qual Life Res*. 2005;14:207–219.
10. Lamers LM, Bouwmans CAM, van Straten A, et al. Comparison of EQ-5D and SF-6D utilities in mental health patients. *Health Econ*. 2006;15:1229–1236.
11. Barton GR, Bankart J, Davis AC, et al. Comparing utility scores before and after hearing-aid provision: results according to the EQ-5D, HUI3 and SF-6D. *Appl Health Econ Health Policy*. 2004;3:103–105.
12. Harrison MJ, Davies LM, Bansback NJ, et al. The validity and responsiveness of generic utility measures in rheumatoid arthritis: a review. *J Rheumatol*. 2008;35:592–602.
13. Grieve R, Grishchenko M, Cairns J. SF-6D versus EQ-5D: reasons for differences in utility scores and impact on reported cost-utility. *Eur J Health Econ*. 2009;10:15–13.
14. Barton GR, Sach TH, Avery AJ, et al. Lifestyle interventions for knee pain in overweight and obese adults aged ≥45 years: cost-effectiveness analysis. *BMJ.* In press.
15. HM Treasury. *The Green Book Appraisal and Evaluation in Central Government*. London, UK: Stationery Office; 2003.
16. Schafer JL. Multiple imputation: a primer. *Stat Methods Med Res*. 1999;8:3–15.
17. Dolan P. Modelling valuations for EuroQol health states. *Med Care*. 1997;35:1095–1108.
18. Ware JE, Sherbourne C. The MOS 36 item short-form health survey: conceptual framework and item selection. *Med Care*. 1992;30:473–483.
19. Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. *Health Econ*. 2005;14:487–496.
20. Briggs AH, Clark T, Wolstenholme J, et al. Missing . . . presumed at random: cost-analysis of incomplete data. *Health Econ*. 2003;12:377–392.
21. Cantor SB. Cost-effectiveness analysis, extended dominance, and ethics: a quantitative assessment. *Med Decis Making*. 1994;14:259–265.
22. National Institute of Health and Clinical Excellence. Guide to the Methods of Technology Appraisal. NICE Publications; 2008.
23. Barton GR, Briggs AH, Fenwick EA. Optimal cost-effective decisions: the role of the cost-effectiveness acceptability curve (CEAC), cost-

effectiveness acceptability frontier (CEAF) and expected value of perfect information (EVPI). *Value Health*. 2008;11:886–897.
24. Fenwick E, Claxton K, Sculpher MJ. Representing uncertainty: the role of cost-effectiveness acceptability curves. *Health Econ*. 2001;10:779–787.
25. Briggs AH, Sculpher MJ, Claxton K. Decision Modelling for Health Economic Evaluation. New York, NY: Oxford University Press; 2006.
26. Claxton K. Bayesian approaches to the value of information: implications for the regulation of new pharmaceuticals. *Health Econ*. 1999;8:269–274.
27. Briggs AH, Wonderling DE, Mooney CZ. Pulling cost-effectiveness analysis up by its bootstraps: a non-parametric approach to confidence interval estimation. *Health Econ*. 1997;6:327–340.
28. Stinnett AA, Mullahy J. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. *Med Decis Making*. 1998;18:S68–S80.
29. Claxton K, Sculpher M, Drummond M. A rational framework for decision making by the National Institute For Clinical Excellence (NICE). *Lancet*. 2002;360:711–715.
30. Torrance GW. Measurement of health state utilities for economic appraisal. *J Health Econ*. 1986;5:1–30.
31. Dolan P, Gudex C, Kind P, et al. Valuing health states: a comparison of methods. *J Health Econ*. 1996;15:209–231.
32. Barton GR, Sach TH, Avery AJ, et al. A comparison of the performance of the EQ-5D and SF-6D for patients aged ≥45 years. *Health Econ*. 2008;17:815–832.
33. Michaels JA, Brazier JE, Campbell WB, et al. Randomized clinical trial comparing surgery with conservative treatment for uncomplicated varicose veins. *Br J Surg*. 2006;93:175–181.
34. van Stel HF, Buskens E. Comparison of the SF-6D and the EQ-5D in patients with coronary heart disease. *Health Qual Life Outcomes*. 2006;4:20.
35. Thomas KJ, MacPherson H, Ratcliffe J, et al. Longer term clinical and economic benefits of offering acupuncture care to patients with chronic low back pain. *Health Technol Assess*. 2005;9:1–109.
36. Fitzpatrick R, Davey C, Buxton MJ, et al. Criteria for assessing patient based outcome measures for use in clinical trials. *Health Technol Assess*. 1998;14:1–74.
37. Brazier JE, Deverill M, Green C, et al. A review of the use of health status measures in economic evaluation. *Health Technol Assess*. 1999;3:1–164.
38. Pickard AS, Wang Z, Walton SM, et al. Are decisions using cost-utility analyses robust to choice of SF-36/SF-12 preference-based algorithm? *Health Qual Life Outcomes*. 2005;3:11.
39. Michaels JA, Campbell WB, Brazier JE, et al. Randomised clinical trial, observational study and assessment of cost-effectiveness of the treatment of varicose veins (REACTIV trial). *Health Technol Assess*. 2006;10:1–196.
40. van den Hout WB, de Jong Z, Munneke M, et al. Cost-utility and cost-effectiveness analyses of a long-term, high-intensity exercise program compared with conventional physical therapy in patients with rheumatoid arthritis. *Arthritis Rheum*. 2005;53:39–47.
41. Marra CA, Marion SA, Guh DP, et al. Not all "quality-adjusted life years" are equal. *J Clin Epidemiol*. 2007;60:616–624.
42. Ware JE, Kosinski M, Keller SD. A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. *Med Care*. 1996;34:220–233.
43. Brazier JE, Ratcliffe J, Salomon JA, et al. *Measuring and Valuing Health Benefits for Economic Evaluation*. New York, NY: Oxford University Press Inc; 2007.