Measuring the level of uncertainty:

The impact of neglecting to include all options in a cost-effectiveness analysis

HESG paper number 29

Garry R. Barton,[1] Tracey H. Sach,[1,2], Elisabeth A. L. Fenwick[3]

[1] Health Economics Group, School of Medicine, Health Policy and Practice, University of East Anglia, Norwich, UK.

[2] School of Chemical Sciences and Pharmacy, University of East Anglia, Norwich, UK

[3] Public Health & Health Policy, University of Glasgow, Glasgow, UK

Contact e-mails: g.barton@uea.ac.uk; t.sach@uea.ac.uk; e.fenwick@clinmed.gla.ac.uk

Objective: To estimate how measures of uncertainty can change when one neglects to include all options in a cost-effectiveness analysis.

Methods: Four options were compared within a randomised controlled trial. Over the 2 year trial period the change in cost to health-service was estimated, along with the QALY gain. In the base-case analysis, levels of cost-utility were calculated along with the cost-effectiveness acceptability curve (CEAC), for each option, and the expected value of perfect information (EVPI). Sensitivity analyses were conducted, including assessing the impact that excluding A) dominated options and B) both dominated and extendedly dominated options had on the EVPI and CEACs for remaining options.

Results: 389 participants were recruited. Option 1 had an estimated incremental cost per QALY of £10,469 (compared to option 4), option 2 was dominated by option 1, and option 3 was subject to extended dominance. In the base-case, at a threshold of λ=£20,000/QALY, the probability of being cost-effective (CEAC level) was between 23.1% and 26.0% for options 1 to 4. The EVPI was £7548 per patient. In sensitivity analysis A the probabilities were 35.0%, 33.3% and 31.8% for options 1, 3, and 4, respectively. The EVPI was £5639. In sensitivity analysis B the probabilities were 50.9% and 49.1% for options 1 and 4, respectively. The EVPI was £3341.

Conclusion: The probability of making the wrong decision and the value of further research can be underestimated when options are excluded from an evaluation, even if they are dominated or subject to extended dominance.

INTRODUCTION

Economic evaluations are undertaken in order to inform decisions about how to allocate scarce health care resources [1,2]. Within such evaluations it has been argued that two key questions should be addressed [3-6]. First, which option is estimated to be cost-effective, on the basis of existing evidence? Second, should further research be undertaken in order to reduce the level of uncertainty associated with that decision? When evaluating potential interventions for a particular population group it is recommended that all relevant alternatives are evaluated [1]. However, a number of recent papers have not included all options in each analysis that is undertaken. For example, when evaluating the treatment of hormone-refractory metastatic prostate cancer researchers produced two models in which they evaluated 3 and 8 options, respectively [7]. Similarly, rather than comparing all options simultaneously, others have undertaken binary comparisons. For example, Garside et al. [8] evaluated two different types of chemotherapy (BCNU-W and TMZ) along with current standard treatment with surgery and radiotherapy, but when estimating the probability of being cost-effective each of the different types of chemotherapy were separately compared with surgery and radiotherapy, they were not directly compared to each other.

In light of the above, within this paper, we seek to assess the impact that reducing the number of options considered when addressing the first question might have on the second aforementioned question – the estimated level of uncertainty associated with a decision. We focus on two measures that provide estimates of the level of uncertainty – the cost-effectiveness acceptability curve (CEAC) [9] and expected value of perfect information (EVPI) [10] – and assess how these measures change when fewer options are evaluated. The analysis is performed using a previously published data set [11] which evaluated four options for people who were overweight and had self-reported knee pain. We re-analyse the data set when certain options are excluded and assess the impact that this has on the CEAC and the EVPI. These re-analyses are akin to what might have happen if a randomised controlled trial (RCT) had been undertaken without including all possible options. As many RCTs often do not include all possible options, e.g. Elliott et al. [12] found that few head-to-head trials of different NSAIDs are undertaken, these re-analyses can provide an assessment of whether it is appropriate to estimate the level of uncertainty using only RCT data which may not include all possible options for the population group in question.

METHODS

Participants

All participants were taking part in the Lifestyle Interventions for Knee Pain (LIKP) study – a randomised controlled trial which was designed to compare the effectiveness and cost-effectiveness of four different options i) diet and strengthening exercise advice, ii) dietary advice, iii) strengthening exercise advice, iv) leaflet provision. Leaflet provision represents usual practice, where hereafter these options are referred to as 1, 2, 3 and 4 as the main focus of this paper is methodological. All registered patients in five 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 four 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 detailed elsewhere [11], briefly they were as follows. Three cost components were estimated – costs associated with i) visits by the health care professionals providing each option, ii) provision of exercise bands to participants in either of the strengthening exercise groups (options 1 and 3), and iii) the change in the cost of analgesics. All costs were estimated over the two year trial period at 2006/7 price levels, costs incurred in the second year were discounted a rate of 3.5% [13], and multiple imputation was used where data was missing [14].

Estimating QALYs

Participants were asked to complete the EQ-5D at baseline, 6, 12, and 24 months post-randomisation. When completing 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 [15]. Up to 243 different health state descriptions can be derived from responses to these five dimensions, and utility scores were assigned to each reported state using the York A1 tariff [16]. After using multiple imputation [14] to predict missing EQ-5D scores, the area under the curve method (with adjustment for baseline differences) [17] was used to estimate the quality adjusted life year

(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%).

Base-case analysis

Cost-utility

Estimates of cost-utility were calculated from the perspective of the UK health service. Dominated options (those with both a higher mean change in cost and lower mean QALY gain, compared to 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) [18]. Subsequently, the cost-utility of non excluded 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 Institute of Health and Clinical Excellence (NICE) [19], we sought to identify the most cost-effective option at the threshold (λ) of £20,000 per QALY.

Decision Uncertainty

In order to depict the level of uncertainty associated with the decision as to which option was cost-effective we constructed the cost-effectiveness acceptability curve (CEAC) [6,20] and the expected value of perfect information (EVPI) [10,21]. The CEAC depicts the probability that an option is cost-effective at different levels of λ (i.e. according to how much one is willing to pay for a QALY gain). It was constructed using non-parametric bootstrapping [22], with 10,000 iterations. The probability of being cost-effective was then equivalent to the proportion of the 10,000 iterations for which each option had the highest net benefit [23] (net monetary benefit= λ*E – C, where E is the QALY gain and C is the change in cost). The CEAC was constructed from the proportions for a range of values of λ. Specific probabilities were reported for a λ of £20,000 per QALY.

In order to estimate the value associated with the level of uncertainty the EVPI was estimated, where the EVPI seeks to estimate the upper bound of the potential value of undertaking further research.[10] In line with previous studies [6], we estimated the EVPI (per patient) by first calculating the value of perfect information in each of the aforementioned iterations (the difference between the highest net benefit across each of the four options and the

net benefit of the most cost-effective option) and then taking the expectation of the value of perfect information across each simulation. Specific values were also reported for a λ of £20,000 per QALY.

Sensitivity analysis

For the reasons outlined in the Introduction, we sought to estimate what would have happened to estimates of uncertainty had the LIKP study been undertaken with fewer than four options. In the first sensitivity analysis (A) we estimated values for the CEAC and EVPI when dominated options were excluded – this was justified on the basis that if the LIKP study (or another RCT) were to have been conducted with fewer options this might have been the most likely option to have been excluded. In sensitivity analysis B, for similar reasons, we re-estimated values for the CEAC and EVPI when dominated options and those that were subject to extended dominance were excluded from the analysis. Finally, in sensitivity analysis C, we re-analysed results comparing just the two most effective options. Here, option 4 (leaflet provision), which might be similar to a placebo in a drug trial, was thereby excluded in order for this reanalysis to be akin to a two option head-to-head RCT.

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.

Costs

Over the 2 year trial period the mean overall change in cost to the health service for those who received option 1 was estimated to be £615.64, compared to £735.57 for option 2, £214.66 for option 3 and –£31.07 for option 4 (leaflet provision) (see Table 1).

<< Table 1 >>

QALYs

The EQ-5D was completed, on average, at baseline, 6, 12 and 24 months by 86.5% of participants. After imputing missing EQ-5D scores, the mean QALY gain was estimated for each

of the four options (see Table 1). It can be seen that leaflet provision was associated with the lowest gain (0.085 QALYs) and that dietary intervention plus quadriceps strengthening exercises was associated with the largest gain (0.147 QALYs).

Base-case analysis

Cost-utility

By comparing the mean change in cost and mean QALY gain across each option (see Table 1) it was apparent that option 2 was dominated by option 1. Similarly, option 3 was not cost-effective as it was subject to extended dominance (combinations of option 4 and option 1 could provide a higher benefit at equivalent cost). Compared to option 4, option 1 had a mean incremental cost of £646.71 and a mean incremental QALY gain of 0.062, giving an ICER of £10,469.44. Thus, at a λ of £20,000 per QALY option 1 was estimated to be cost-effective.

Decision Uncertainty

The CEACs for each of the four options are plotted in Figure 1. It can be seen that at low levels of λ (<£5,000 per QALY) option 4 was estimated to have the highest probability of being cost-effective. However, for values >£5,000 per QALY the probability of each option being cost-effective was <30%. For a value of λ=£20,000 per QALY the probability of each option being cost-effective, as estimated by the individual CEACs, are illustrated in Table 1. Option 1 had a 26.0% probability, which meant that in choosing option 1 it was estimated that there was a 74.0% chance of making the wrong decision. This suggests that there is much uncertainty associated with the decision as to which option is most cost-effective. The EVPI is plotted in Figure 1, at a λ of £20,000 per QALY it was estimated to equate to £7,195 per patient (see Table 2).

<< Figure 1 >>

<< Table 2 >>

Sensitivity analysis

In the base-case analysis option 2 was dominated by option 1. Thus, in sensitivity analysis A we re-estimated values for the CEAC and EVPI when option 2 was excluded from the analysis. According to the CEAC, the probability of being cost-effective increased for all three remaining options – to 35.0% for option 1, 33.3% for option 3 and 31.8% for option 4 (see Table 1 and

Figure 2). The incremental cost per QALY values remained unchanged (option 1 was still estimated to be cost-effective), but by choosing option 1, when only three options were considered, it would have been estimated that the probability of making the wrong decision was 65.0% at a value of λ=£20,000 per QALY. Similarly, the value of undertaking further research, as estimated by the EVPI, fell and was estimated to be £5,639 at a value of λ=£20,000 per QALY (see Figure 2).

In the base-case analysis option 3 was subject to extended dominance. Thus, in sensitivity analysis B we re-estimated values for the CEAC and EVPI when both options 2 and 3 were excluded from the analysis. According to the CEAC the probability of the remaining options (1 and 4) being cost-effective increased beyond that in both the base-case and sensitivity analysis A (see Table 1 and Figure 2). Indeed by choosing option 1, when only these two options were evaluated, the probability of making the wrong decision would have been estimated to be lower at 49.1% for a value of λ=£20,000 per QALY. The value of undertaking further research would also have been estimated to be lower, at £3,341 per patient (see Table 2).

In the base-case analysis options 1 and 2 were estimated to be most effective so sensitivity analysis C included only these two options, akin to a head-to-head RCT with no placebo arm. Had this binary comparison been undertaken, option 1 would have been estimated to have a 52.6% chance of being cost-effective at a value of λ=£20,000 per QALY, compared to 48.4% for option 2 i.e. the probability of making the wrong decision would have been estimated to be 48.4% by choosing option 1. Interestingly, in contrast to the base-case analysis, option 1 was here estimated to have a higher probability of being cost-effective (at a λ of £20,000 per QALY) than option 2 (see Table 1), demonstrating that excluding options can result in a reversal of the rank ordering of the probability of cost-effectiveness according to the CEAC. At the same level of λ the value of further research, as estimated by the EVPI, would also have been estimated to be lower than in the base-case at a value of £3,854 per patient (see Table 2).

DISCUSSION

In this study we have shown that excluding options, even those that are not cost-effective, led to increased estimates of the probability of remaining options being cost-effective (CEAC values). In contrast, estimates as to the value of further research were lower. Additionally, in sensitivity analysis C the rank ordering of options (from highest to lowest probability of being

cost-effective) differed from that in the base-case analysis.

Explanations

Removing an option works in the same way removing a political party might work after the first round of an election as, in the second round, the votes for the excluded party (option) are transferred to another party (option). Moreover, in the same way that in the second round the proportion of votes received does not equate to the percentage of the population which would prefer that party to be in power, when options are excluded the CEAC will increase to a level which is higher than the true probability that an option is cost-effective compared to all other options.

The EVPI is determined by both the probability, and consequences, of making the wrong decision. As explained above, as the estimated probability of being cost-effective increases when there are fewer options the probability of making the wrong decision falls. Similarly, the consequences of making the wrong decision are determined by the difference in net benefit between options, and, in the analyses conducted in this paper, the exclusion of options resulted in the difference in net benefit (across options) being, on average, lower than when all options were included.

Comparisons with other studies

In line with our results, the aforementioned study by Collins et al. [7] found that when only three options were evaluated the CEACs were higher with levels of 39%, 39% and 22%, respectively, at a value of λ=£20,000 per QALY. In contrast, the same options were estimated to have probabilities of 28%, 18% and 7%, respectively, when all 8 options were evaluated.

Garside et al. [8] evaluated two different types of chemotherapy (BCNU-W and TMZ) along with current standard treatment with surgery and radiotherapy (usual care), however when estimating the CEACs each type of chemotherapy was separately compared to usual care. Thus it was stated that, compared with usual care, (for an λ of £30,000 per QALY) there was an 11% probability that BCNU-W was cost-effective and a 23% probability that TMZ was the most cost-effective option [8] i.e. usual care had a probability of 89% probability in the first comparison and 77% in the second comparison. Our analyses have hopefully demonstrated that had Garside et al. [8] evaluated all three options simultaneously one would have expected all of the

aforementioned probabilities to have been lower. i.e. by undertaking binary comparisons the probability of options being cost-effective can only be overestimated.

Limitations

We have considered how estimates as to the level of uncertainty would have varied had fewer options been included in our evaluation. The main limitation of our study is however that, though there are implications for other studies (see below), the impact of excluding options is likely to be study specific. This is due to the fact that it has previously been demonstrated that the shape of both the CEAC and EVPI are dependent on both the correlation structure and the differences in the level of variation between multiple options [6], both of which are likely to vary across studies. Indeed, though we would argue that excluding options can only increase the probability that remaining options are deemed most cost-effective (one exception would be if an option which never had the highest net benefit in any iteration was excluded, here CEACs would remain unchanged), the effect on the EVPI is less certain. We found that the EVPI always decreased when options were excluded, however the EVPI is determined by the difference between the highest net benefit across each option and the net benefit of the most cost-effective option. Thus, if one were (for some reason) to exclude the most cost-effective option this would entirely change the comparator in the EVPI analysis, which could result in either a higher or lower difference in net benefit and thereby an increase or decrease in the value of further research (compared to the situation where all options were evaluated).

Implications

It has been argued that the CEACs are important to policy makers as they represent an appropriate method of estimating the probability of making an error [24]. However, it has also been pointed out that CEACs may be misinterpreted [25,26]. CEACs were designed to show the probability that a treatment option is cost-effective (at different levels of λ) [9,24], but when options are excluded we have shown that this may not be depicted. Instead, the probability of cost-effectiveness will only have been estimated in relation to the (smaller) number of options which have been evaluated. In spite of the fact that authors might be careful to clarify that these are e.g. only binary CEACs, and do not include all options, we would argue that such analyses could lead to further misinterpretations. A further implication is that when a new technology

(option) becomes available, even if it is not estimated to be cost-effective (on the basis of existing evidence), it may have an impact on the level of uncertainty, both in terms of the estimated probability of cost-effectiveness (CEAC) and the value of further research (EVPI).

CONCLUSION

By constructing CEACs when not all options are evaluated the probability of remaining options being cost-effective can be overestimated. As a result the probability of making the wrong decision will be underestimated, as can the consequences of making the wrong decision, which would mean that the EVPI (value of further research) will also be underestimated. Thus by excluding options, even those that are subject to extended dominance or dominated, the level of uncertainty might be underestimated.

REFERENCES

1. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the Economic Evaluation of Health Care Programmes (3rd Edition). New York: Oxford University Press; 2005.
2. Sach TH, Barton GR, Doherty M, Muir K, Jenkinson C, Avery AJ. The relationship between BMI and health related quality of life: comparing the EQ-5D, EuroQol VAS, and SF-6D. Int J Obes Relat Metab Disord 2007; 31: 189-196.
3. Claxton K, Sculpher M, Drummond M. A rational framework for decision making by the National Institute For Clinical Excellence (NICE). Lancet 2002; 360: 711-5.
4. Sculpher M, Claxton K. Establishing the cost-effectiveness of new pharmaceuticals under conditions of uncertainty: when is there sufficient evidence? Value Health 2005; 8: 433-446.
5. Fenwick E, Palmer S, Claxton K, Sculpher M, Abrams K, Sutton A. An iterative Bayesian approach to health technology assessment: Application to a policy of pre-operative optimization for patients undergoing major elective surgery. Med Decis Making 2006; 26: 480-496.
6. 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.

7. Collins R, Fenwick EA, Trowman R, Perard R, Norman G, Light K et al. A systematic review and economic model of the effectiveness and cost-effectiveness of docetaxel in combination with prednisone or prednisolone for the treatment of hormone-refractory metastatic prostate cancer. CRD/CHE Technology Assessment Group: University of York; 2005.
8. Garside R, Pitt M, Anderson R, Rogers G, Dyer M, Mealing S et al. The effectiveness and cost-effectiveness of carmustine implants and temozolomide for the treatment of newly diagnosed high-grade glioma: a systematic review and economic evaluation. 2007; 11 (45): 1-242.
9. van Hout BA, Al MJ, Gordon GS, Rutten FF. Costs, effects and C/E-ratios alongside a clinical trial. Health Econ 1994; 3: 309-319.
10. Claxton K. Bayesian approaches to the value of information: implications for the regulation of new pharmaceuticals. Health Econ 1999; 8: 269-274.
11. Barton GR, Sach TH, Avery AJ, Doherty M, Jenkinson C, Muir KR. Lifestyle Interventions for Knee Pain in overweight and obese adults aged ≥45 years: Cost-effectiveness analysis. BMJ in press;
12. Elliott RA, Hooper L, Payne K, Brown TJ, Roberts C, Symmons D. Preventing non-steroidal anti-inflammatory drug-induced gastrointestinal toxicity: are older strategies more cost-effective in the general population? Rheumatology (Oxford) 2006; 45: 606-613.
13. HM Treasury. The green book appraisal and evaluation in central government. London: Stationery Office; 2003.
14. Schafer JL. Multiple imputation: a primer. Stat Methods Med Res 1999; 8: 3-15.
15. Brooks R. EuroQol: the current state of play. Health Policy 1996; 37: 53-72.
16. Dolan P. Modelling valuations for EuroQol health states. Med Care 1997; 35: 1095-1108.
17. Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. Health Economics 2005; 14: 487-496.

18. Cantor SB. Cost-effectiveness analysis, extended dominance, and ethics: a quantitative assessment. Med Decis Making 1994; 14: 259-265.
19. National Institute of Health and Clinical Excellence. Guide to the Methods of Technology Appraisal. NICE publications; 2008.
20. Fenwick E, Claxton K, Sculpher MJ. Representing uncertainty: the role of cost-effectiveness acceptability curves. Health Econ 2001; 10: 779-787.
21. Briggs AH, Sculpher MJ, Claxton K. Decision modelling for health economic evaluation. New York: Oxford University Press; 2006.
22. 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.
23. 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.
24. Fenwick EA, Briggs AH. Cost-Effectiveness Acceptability Curves in the Dock: Case Not Proven? Med Decis Making 2007; 27: 93-94.
25. Coast J. Is economic evaluation in touch with society's health values? BMJ 2004; 329: 1233-1236.
26. Groot Koerkamp B, Hunink M, Stijnen T, Hammitt JK, Kuntz KM, Weinstein MC. Limitations of acceptability curves for presenting uncertainty in cost-effectiveness analysis. Med Descis Making 2007; 27: 101-111.

Table 1. Estimates of the mean change in cost and mean QALY gain associated with each option, along with overall estimates of cost-effectiveness and the level of uncertainty.

	Option			
	Dietary intervention plus strengthening exercises (1)	Dietary intervention (2)	Strengthening exercises (3)	Leaflet provision (4)
Overall change in cost	£615.64	£735.57	£214.66	-£31.07
QALY gain	0.147	0.133	0.090	0.085
Incremental cost-effectiveness ratio (ICER)	£10,469	Dominated	ED	–
CEAC value (λ =£20,000/QALY) Base-case	26.0%	26.1%	24.8%	23.1%
Sensitivity Analysis A	35.0%	–	33.3%	31.8%
Sensitivity Analysis B	50.9%	–	–	49.1%
Sensitivity Analysis C	51.6%	48.4%	–	–

CEAC = cost-effectiveness acceptability curve

Table 2. Estimates as to how the expected value of perfect information (EVPI) varies, at a value of λ =£20,000/QALY, when different numbers of options are evaluated.

	EVPI (per patient)
Base-case	£7,195
Sensitivity Analysis A	£5,639
Sensitivity Analysis B	£3,341
Sensitivity Analysis C	£3,854

Figure 1. Results of the base-case cost-effectiveness analysis displayed in terms of the cost-effectiveness acceptability curves, and the expected value of perfect information.

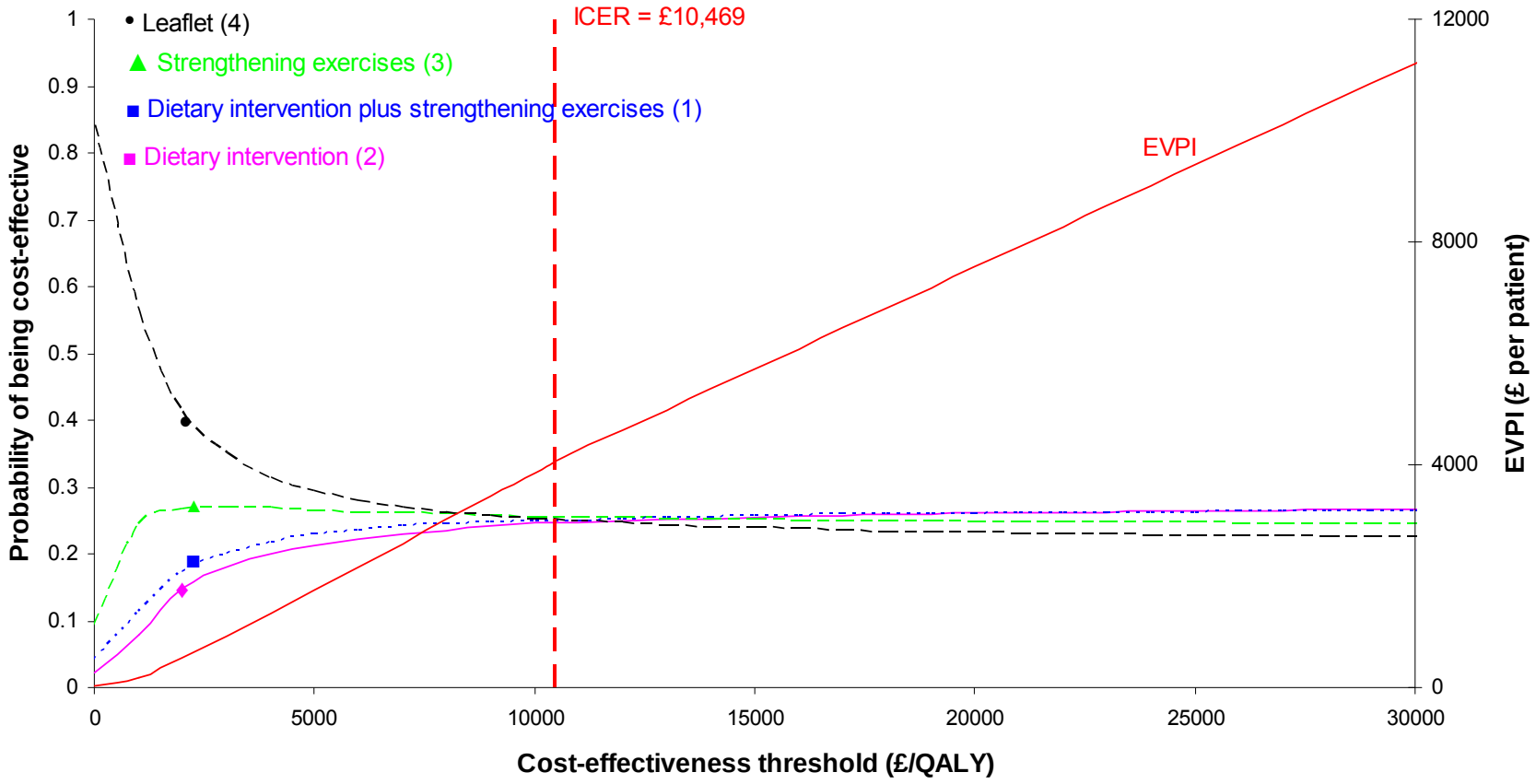

Figure 2. Results of the sensitivity analysis A displayed in terms of the cost-effectiveness acceptability curves, and the expected value of perfect information.

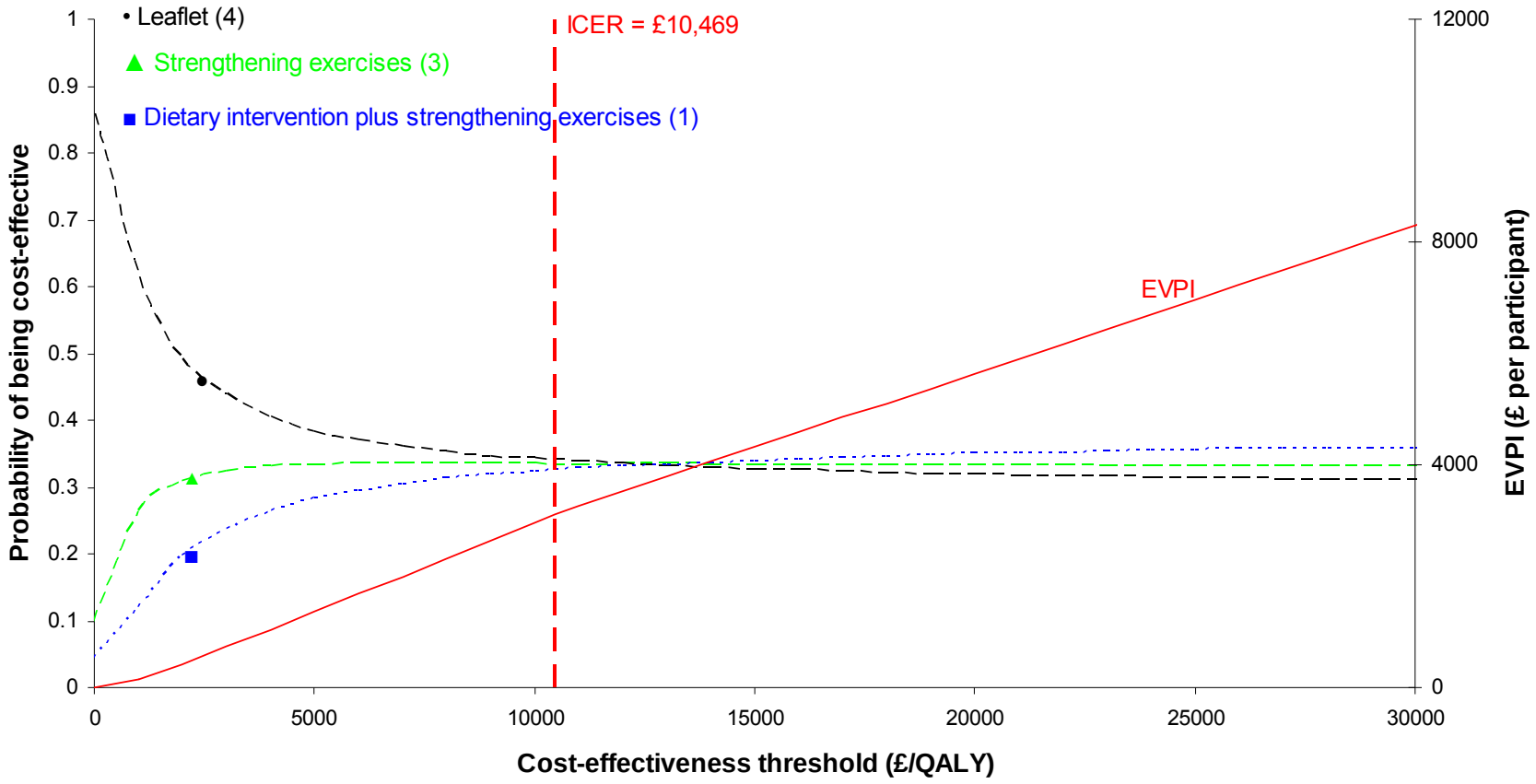

Figure 3. Results of the sensitivity analysis B displayed in terms of the cost-effectiveness acceptability curves, and the expected value of perfect information.

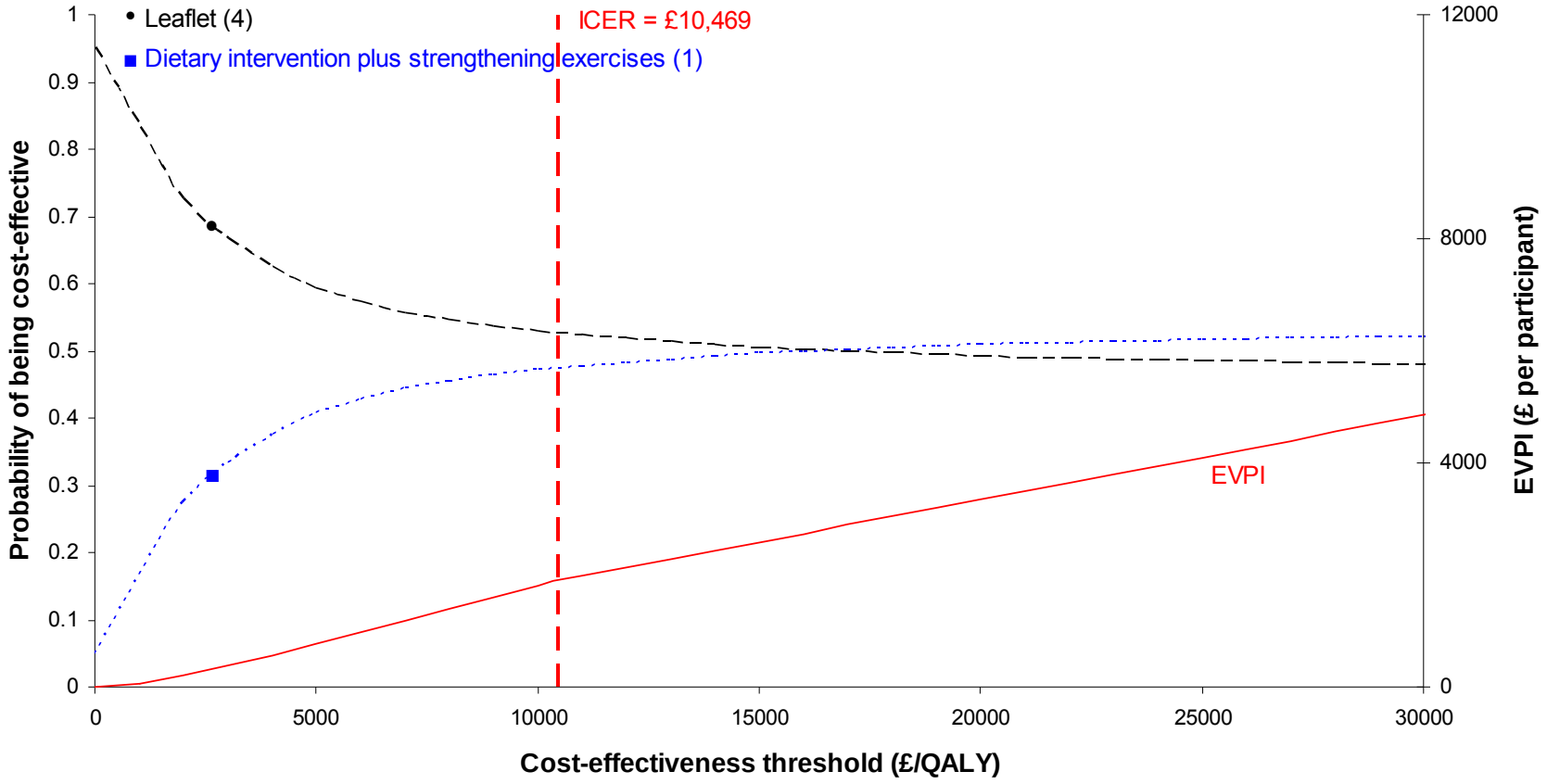

Figure 4. Results of the sensitivity analysis C displayed in terms of the cost-effectiveness acceptability curves, and the expected value of perfect information.