Work in progress: comments welcome

Bayesian cost-effectiveness analysis for frequentists: interpreting cost-effectiveness acceptability curves

Andrew H. Briggs*
Health Economics Research Centre & Nuffield College
University of Oxford, UK.

Abstract
The aim of this paper is to discuss the use of Bayesian methods in cost-effectiveness analysis. Historically, frequentists have been cautious of Bayesian methodology, which is often held as synonymous with a subjective approach to statistical analysis. In this paper, the potential overlap between Bayesian and frequentist approaches to cost-effectiveness analysis are explored – the focus being on the empirical and uninformative prior-based approaches to Bayesian methods rather than the use of subjective beliefs. This approach emphasises the advantage of Bayesian interpretation for decision making while retaining the robustness of the frequentist approach. In particular the use of cost-effectiveness acceptability curves is examined. A traditional frequentist approach is equivalent to a Bayesian approach assuming no prior information, while where there is pre-existing information available from which to construct a prior distribution, an empirical Bayes approach is equivalent to a frequentist approach based on pooling the available data. Cost-effectiveness acceptability curves directly address the decision-making problem in cost-effectiveness analysis. Although it is argued that their interpretation as the probability that an intervention is cost-effective given the data requires a Bayesian interpretation, this should generate no misgivings for the frequentist.

Acknowledgements: This paper has benefited greatly from Dr Alastair Gray's comments and from discussion with Sir David Cox. Responsibility for all remaining errors is of course my own. Funding comes from a joint UK MRC/Anglia and Oxford Region Training Fellowship in Health Economics and Medical Statistics.

---

* Address for correspondence: Andrew Briggs, Health Economics Research Centre, Institute of Health Sciences, University of Oxford, Headington, Oxford OX3 7LF, UK. Tel: +44 1865 226 847. Fax: +44 1865 226 842. Email: andrew.briggs@ihs.ox.ac.uk.

1. Introduction

Economic evaluation of health care interventions is fundamentally concerned with decision-making. In the context of a budget constrained system of health care, the effectiveness of a health care intervention is a necessary but not sufficient condition for provision of that intervention. The costs of health care must also be considered in order to achieve maximum health gain from limited resources. Where a health care intervention is both more costly and more effective, the ratio of the additional cost to additional effect (the incremental cost-effectiveness ratio or ICER) is calculated in order to summarise the cost-effectiveness of the intervention. If this ICER is less than some maximum willingness to pay for additional health gain (or 'ceiling cost-effectiveness ratio', R_c) then the intervention is said to represent 'good value for money'. If, however, the ICER is greater than R_c, the intervention is not cost-effective and resources should be prioritised to other more worthwhile interventions.

Of course, estimates of the cost-effectiveness of health care interventions are subject to uncertainty, which should be taken into account during the decision-making process. With the growth of economic evaluation conducted alongside clinical trials, analysts are increasingly in a position to handle statistically the uncertainty due to sampling variation. This has engendered a growing interest in the use of statistical methods for cost-effectiveness analysis. A recent area of interest has been the calculation of confidence intervals for cost-effectiveness ratios, and more generally the presentation of uncertainty in the results of cost-effectiveness analysis (O'Brien et al. 1994; van Hout et al. 1994; Wakker and Klaassen, 1995; Willan and

O'Brien, 1996; Chaudhary and Stearns, 1996; Briggs et al. 1997; Obenchain et al. 1997; Polsky et al. 1997; Briggs and Fenn, 1998).

The aim of this paper is to explore how Bayesian methods might be used to overcome some of the problems encountered in cost-effectiveness studies when adopting a strictly frequentist approach to handling uncertainty. It is sometimes thought that Bayesians and frequentists inhabit different scientific paradigms that exclude any middle ground. However, there is an increasing acceptance of the fact that it is possible to exploit the natural interpretation associated with a Bayesian approach to statistical analysis while retaining the robustness of the frequentist approach (Carlin and Louis, 1996).

Bayesian methods might usefully be classified into three main types, dependent on the approach to prior information. Empirical Bayes describes the approach of estimating prior distributions on the basis of previously available statistical information. A second approach to Bayes would be to assume an uninformative or reference prior that purports to contain no information concerning the parameter of interest. The third approach could be described as Subjective Bayes, where prior information is elicited (in a coherent fashion) from experts on the basis of their personal beliefs. The purpose of such a classification is to emphasise that only the third type of approach is synonymous with the subjective approach commonly considered (by frequentists) as introducing a lack of robustness to statistical analysis. Hence, in this paper, the emphasis is on the overlap between the frequentist approach and Bayes methods employing empirical or uninformative priors, rather than on the subjective Bayes methods.

The next section introduces the standard frequentist approach to handling uncertainty in economic evaluation and outlines the sort

of problems that can arise. The representation of uncertainty as a cost-effectiveness acceptability curve is argued to provide the sort of information most useful for decision-makers compared to standard confidence intervals. However, the interpretation of such curves is most natural using a Bayesian approach. The third section therefore considers a Bayesian approach to cost-effectiveness analysis. The net-benefits statistic is employed in preference to the ICER due its desirable statistical properties. However, the overall equivalence of net-benefits and the ICER is demonstrated by the equivalence of cost-effectiveness acceptability curves based on the two approaches. The final section offers a discussion of the issues raised in this paper. Illustrative data are employed to elucidate the methods and the focus is on the intuitive appeal of the Bayesian approach rather than the technical aspects: statistical formulae are therefore confined to the appendix.

2. Handling uncertainty in cost-effectiveness analysis

In a trial situation, on the basis of data collected from two groups of patients receiving the alternative therapies, the ICER can be estimated by

$$\hat{R} = \frac{\overline{C}_T - \overline{C}_C}{\overline{E}_T - \overline{E}_C} = \frac{\Delta\overline{C}}{\Delta\overline{E}} \qquad (1)$$

where $\overline{C}_T$ and $\overline{C}_C$ are the mean costs in the treatment and control arms of the trial respectively and $\overline{E}_T$ and $\overline{E}_C$ are the mean effects. A traditional approach for handling uncertainty due to sampling variation would be to estimate the confidence interval for the ICER and compare the interval to the ceiling cost-effectiveness ratio, R_c. If the estimated interval is found to be above R_c then the intervention is

significantly cost-ineffective with the implication that the intervention should not be funded. If the estimated interval is found to be below R_c then the intervention is significantly cost-effective with the implication that the treatment should be funded.

Although this traditional approach does not appear contentious, it does present a number of problems that are illustrated by means of an example. Consider a trial-based economic analysis that generated the data shown in Table 1.[1] Note from the table that the new therapy has been shown to generate significant health outcome effects over the existing (control) therapy. In practice, this is likely to mean that there will be pressure to implement this intervention for the relevant patient group. However, the results also show the intervention to be significantly more costly than the existing treatment for these patients. Consideration should be given therefore to the ICER, which from Table 1 is estimated to be £4,836 per unit of health gain. Suppose we know that the appropriate ceiling ratio for this decision is £10,000 per unit effect, then the intervention looks as if it may offer good value for money – if the confidence interval for the ICER excludes this value then the intervention should be implemented.

Ratio statistics pose particular problems for standard methods of calculating confidence intervals when there is a non-negligible probability that the denominator of the ratio can take a very small value. Applying the non-parametric bootstrap method (Chaudhary and Stearns, 1996; Obenchain et al. 1997; Briggs et al. 1997) allows

---

[1] Although these are hypothetical data, they are illustrative of a 'typical' economic analysis alongside a clinical trial. Real data could have been used at this point, which may be seen as preferable to hypothetical data. However, the Bayesian analysis described in the next section makes use of 'existing data' for the same 'intervention'. Without access to such prior data for any real life examples, it seemed preferable to keep the whole example illustrative.

the uncertainty in cost-effectiveness to be 'visualised' and emphasises some of the problems for handling uncertainty in cost-effectiveness analysis. Figure 1 shows the bootstrap estimate of the sampling distribution of the ICER for the data from Table 1, based on 1,000 bootstrap replications of the data in the treatment and control groups. It is immediately apparent that the sampling distribution of the ICER in the case of the data from Table 1 does not follow a well behaved distributional form. In fact, the histogram of Figure 1 is truncated in order to present the distribution more clearly. The most straightforward method of obtaining a bootstrap 95% confidence interval is to select the 2.5th and 97.5th percentile points of the vector of bootstrap replicates of the ICER – equivalent to cutting 2.5% from the tails of the estimated distribution shown in Figure 1. The resulting interval on the basis of the bootstrap replicates from Figure 1 is -£772 to £24,646.

Note, however, that this interval is seriously misleading. Consider Figure 2, which shows the same bootstrap replicates of the cost and effect differences plotted on the cost-effectiveness (CE) plane. It is immediately apparent that although the majority of bootstrap cost/effect pairs are falling in the north-east quadrant of the plane where the ICER is positive, an important minority of cost/effect pairs fall in the south-east and north-west quadrants of the plane where the ICER is negative. In the rank ordering process underlying the histogram of Figure 1 and the percentile based confidence interval, negative ratios are grouped together. However, from a decision-making perspective negative ratios from different quadrants of the CE plane are fundamentally different – a negative ratio due to a negative cost difference implies that the new treatment is both cheaper and more effective than that currently provided, while a negative ratio due

to a negative effect difference implies the exact opposite. A practical solution to this problem would be to simply rank the negative ratios due to negative effects above the highest positive ICER replicate and then estimate the percentile confidence limits. While valid from a decision-making perspective, this approach is rather suspect statistically since the problem identified above is a problem not of the statistical method for calculating confidence limits (which is valid) but a problem of interpreting a one-dimensional ratio, when the decision-making space is in fact two-dimensional (Briggs and Fenn, 1998).

Recently, a net-benefits approach has been suggested that provides a solution to the problem, providing that the ceiling ratio appropriate for decision making, R_c is known (Stinnett and Mullahy, 1998; Tambour et al. 1998). The approach involves using the value of R_c to rescale either the effect difference or the cost difference in order to provide a net benefit statistic on the cost (Tambour et al. 1998) or the effect scale (Stinnett and Mullahy, 1998). This is the same approach adopted by Phelps and Mushlin to argue for the (near) equivalence of cost-benefit and cost-effectiveness analysis, although they did not suggest that the statistic had any desirable properties related to handling uncertainty (Phelps and Mushlin, 1991). Since the net-benefit statistics are exactly equivalent, the choice of scale is a matter of preference. For this paper, the net benefits on the cost scale are used and defined as

$$NB = R_c \cdot \Delta\overline{E} - \Delta\overline{C} . \qquad (2)$$

Positive net benefits for an intervention indicate that the intervention represents good value for money. Therefore, the standard statistical approach would be to estimate the confidence interval for net-benefits and to see whether that interval excludes zero. Note that no problems

of interpretation for the net-benefits statistic arise: negative effect differences with positive cost differences give negative net-benefits, while positive effect differences and negative cost-differences give positive net-benefits. In terms of Figure 2, employing R_c = £10,000 will lead to positive net-benefits for bootstrap replications to the right of the line representing the decision rule and negative net-benefits to the left. A further advantage of the net-benefit statistic is its well behaved statistical properties. Consider Figure 3, showing the histogram of the bootstrap estimate of the sampling distribution of the net-benefit statistic from Equation 2 as compared with the bootstrap estimate of the sampling distribution for the ICER in Figure 1 (note that the same bootstrap replications are used in both figures).

Despite the desirable properties of the net-benefit statistic, two problems remains. First, what are the implications for decision-making when the net-benefit statistic is non-significant? For example, Table 1 gives the bootstrap estimate of the 95% confidence interval for net-benefits as -£3,648 to £11,260 indicating that the intervention is not significantly cost-effective at the 5% level. Stinnett and Mullahy (1998)advocate plotting out a P-value function for net-benefits, thereby alerting the decision-maker to the confidence level at which the statistic just becomes significant. While this is a useful approach, it does not address the second problem that the net-benefit statistics relies on a predefined value for R_c whereas in fact the value of R_c is itself unknown.

Both these problems can be solved simultaneously by plotting the cost-effectiveness acceptability curve (van Hout et al. 1994). The (one-sided) confidence level at which the net-benefits just become significant can be plotted as a function of R_c. This approach was

initially described in relation to cost-effectiveness ratios and the cost-effectiveness plane (van Hout et al. 1994), and can be equivalently thought of as considering the proportion of bootstrap replications falling to the right hand side of the line representing R_c in Figure 2 as the line is rotated anti-clockwise from the horizontal ($R_c = 0$) through to the vertical ($R_c = \infty$). The cost-effectiveness acceptability curve for the data from Table 1 is presented in Figure 4. The curve cuts the vertical axis at 0.022, which corresponds to the (one-sided) P-value for cost difference. The point of inflexion for the curve occurs at the point estimate of the ICER (£4,836) and the curve is tending to the value 0.969, which corresponds to one minus the (one-sided) P-value for the effect difference. If R_c for a particular decision-maker were £10,000, then Figure 4 can be used to determine that such a result would only be significant at the 20% level.

Although a strict frequentist interpretation of cost-effectiveness acceptability curves is possible through the use of P-value as described above, the natural way to interpret these curves (as illustrated by the labelling of the vertical axis in Figure 4) is as the probability that the intervention is cost-effective. Indeed, this is the way cost-effectiveness acceptability curves have been presented in the literature to date (van Hout et al. 1994; Raikou et al. 1998; Briggs and Fenn, 1998). This natural interpretation is exactly analogous to that of a 95% confidence interval as representing a 95% probability that the true parameter is contained within the interval when the frequentist interpretation is that the true parameter will be contained within the interval 95% of the time in repeated application.

3. Bayesian methods for cost-effectiveness analysis

Under a Bayesian interpretation, parameters of interest are ascribed a distribution reflecting our uncertainty concerning the true value of the parameter. Fundamentally, the Bayesian approach includes a learning process whereby beliefs concerning the distributions of parameters (prior distributions) are updated (to posterior distributions), as information becomes available, through the use of Bayes' Theorem. Historically, advocates of the Bayesian approach were seen to inhabit a different scientific paradigm that was at odds with the frequentist paradigm such that frequentists considered Bayes methods as subjective and highly dependent on the prior beliefs employed, while frequentist methods were objective and robust. However, the adoption of such an extreme position would be to reject a set of very powerful methods that may be of import, even for frequentists (Carlin and Louis, 1996). The empirical Bayes methods and Bayesian analysis based on uninformative prior distributions are not subjective and have much to offer the frequentist analyst. It is these methods that form the basis of the illustration below.

The algebra for Bayes theorem is given in the appendix. As stated above, the intuition for the use of a Bayesian approach is that prior information can be used in conjunction with observed data to form a posterior distribution for the parameter of interest that is influenced both by the observed data and the prior information. It is common for prior distributions to be specified in terms of a distribution that is conjugate to the distribution of the observed data for mathematical convenience. Conjugate priors lead to posterior distribution from the same family of distributions (Rice, 1995), which make the calculation of the posterior distribution more straightforward. Although the advent of powerful desktop computers

and appropriate software has meant that this is no longer strictly necessary, there is added convenience when the distributions can be expressed in terms of recognised distributions. In particular, the normal distribution is self-conjugate such that normal data and normal prior lead to a normal posterior distribution (see appendix).

For cost-effectiveness analysis, it is clear that the net-benefits statistic will be much more convenient to handle in a Bayesian analysis than would the ICER statistic. Suppose that prior to the trial that generated the data from the previous section (see Table 1), additional published information for the same intervention had been available. This prior information is summarised in Table 2. Using the net-benefit approach ($R_c = £10{,}000$), Figure 5 presents this prior information and the posterior distribution arising from employing this information together with the data from Table 1. This is the Empirical Bayes approach introduced above. The alternative would be to employ an uninformative prior such that the posterior distribution produced is dominated by the observed data – either because no prior information is available, or because the analyst wishes to discard that information. The posterior distribution based on the uninformative prior is also shown in Figure 5. It is clear that incorporating the prior information reduces the variance of the posterior distribution, but that the point estimate of net-benefit is weighted most heavily toward the existing data. This is because the empirical Bayes approach weights the prior information in relation to its variance compared to the observed data.

Having estimated the prior and posterior distributions for net-benefits using Bayesian methods, these can then be plotted as a function of R_c in order to generate cost-effectiveness acceptability curves. Figure 6 shows the results – as argued above, it is these

curves that give the information that is most relevant to decision-makers. Notice that the cost-effectiveness acceptability curve based on an uninformative prior is the same as the cost-effectiveness acceptability curve from Figure 4 since without prior information on net-benefits, the posterior distribution is dominated by the information from the observed data (see appendix). Therefore, if the decision rule were that $R_c = £10{,}000$ a decision maker could only be 80% sure that the intervention represents a cost-effective use of resources. By contrast, if the prior information from Table 2 is included, the decision-maker can be over 95% sure that the intervention is cost-effective (analagous to the conventional level of significance required in a frequentist analysis).

4. Discussion

The emphasis in this paper has been on the middle ground between Bayesian and frequentist methods for cost-effectiveness analysis. The application of empirical Bayes methods or Bayes methods employing an uninformative prior are largely uncontroversial. Indeed, it has been argued that Bayesian analysis based on uninformative priors is equivalent to a frequentist analysis based on the observed data (while allowing the more natural Bayesian interpretation associated with distributions of parameters). Similarly, empirical Bayes methodology is equivalent to a frequentist approach based on pooling available data. Because of this close relationship with the frequentist approach, some Bayesians would argue that the methods illustrated in this paper are not in fact Bayesian at all since no subjective beliefs are employed. For example, Spiegelhalter and colleagues (1994) argue that while previous results should form the basis of prior distributions

those results should not specify the distribution completely. Referencing Kass & Greenhouse (1989) they go on to argue that to do so would be to treat historical and current data as exchangeable, which is in essence equivalent to simply pooling the results.

One way in which subjective beliefs could be incorporated into the framework described in this paper, that may be acceptable to frequentists, would be to weight prior information in some way. In the present analysis, the two logical extremes are presented. Using data from another study to describe the prior distribution in the Bayesian analysis gives equal weighting to the prior and observed data (while giving more weight to the data with the least variance). On the other hand, using a noninformative prior distribution where data are available from a previous study effectively gives no weight to this prior data. In practice, it may be that data available from previous studies could be used as a basis for the prior distribution, but that less weight could be given to these prior data than to the observed data from the study. Such an approach has been illustrated by Brophy and Joseph (1995) in their Bayesian reanalysis of the GUSTO trial. They looked at the effect on the conclusions reached by the GUSTO investigators (1993) of using information from three previous trials of thrombolytic therapy as a basis for a prior distribution with various weightings attached to the previous data. Their reanalysis suggested that the clinical benefits of tissue-type plasminogen activator over and above those of streptokinase remain uncertain.

In this paper, the focus has been on estimating the posterior distribution of the parameter of interest (net-benefits) and using this to generate cost-effectiveness acceptability curves. A full Bayesian analysis might specify a loss function for decision-making in order to estimate the optimal decision strategy (i.e. that which minimises the

loss function). Claxton and Posnett (1996) have specified such a loss function in terms of R_c which they have employed to determine the optimal size for clinical trials. However, uncertainty concerning the appropriate value for R_c, and the possibility that such a value might vary between decision-makers and over time, may mean that the clear specification of the posterior distribution that could then be used by decision-makers in conjunction with their own loss function may be more appropriate, although this is undoubtedly an area for further research.

At present, and most likely in the future, health economists conducting economic analyses alongside clinical trials will have to work within the sample size constraints imposed by clinical investigators. This is likely to generate the situation where important economic differences cannot be detected at conventional levels of power and significance. The Bayesian approach offers an important and powerful set of methods to the economic analyst seeking to advise policy decisions. This paper has demonstrated how Bayesian methods can be used in a way entirely consistent with the desire by frequentists for an objective approach to statistical analysis. In particular, the Bayesian approach allows a more natural way of interpreting cost-effectiveness acceptability curves. It would be unfortunate if the potential for Bayesian methods were cast aside due to the prior beliefs of frequentists that the methods are necessarily subjective!

Appendix

With sufficient sample size, the distribution of the relevant estimators for the ICER (Equation 1) and net-benefits (Equation 2) can be approximated by normal distributions with the following mean and variance

$$\begin{aligned}
\overline{C}_C &\sim N(\mu_{CC}, \sigma^2_{CC}/n_C) \\
\overline{E}_C &\sim N(\mu_{EC}, \sigma^2_{EC}/n_C) \\
\overline{C}_T &\sim N(\mu_{CT}, \sigma^2_{CT}/n_T) \\
\overline{E}_T &\sim N(\mu_{ET}, \sigma^2_{ET}/n_T)
\end{aligned}$$

where μ and σ^2 represent the true population mean and variance for (subscripts) costs (C) and effects (E) in the treatment (T) and control (C) groups.

Since the sum of two normal variables is itself normally distributed, with sufficient sample size the estimates of the incremental costs and effects are distributed

$$\begin{aligned}
\Delta\overline{E} &\sim N(\mu_{ET} - \mu_{EC}, \sigma^2_{\Delta E}) \\
\Delta\overline{C} &\sim N(\mu_{CT} - \mu_{CC}, \sigma^2_{\Delta C})
\end{aligned}$$

where:[2]

$$\begin{aligned}
\sigma^2_{\Delta E} &= \frac{\sigma^2_{ET}}{n_T} + \frac{\sigma^2_{EC}}{n_C} \\
\sigma^2_{\Delta C} &= \frac{\sigma^2_{CT}}{n_T} + \frac{\sigma^2_{CC}}{n_C}.
\end{aligned} \tag{3}$$

As a ratio of two asymptotically normal variables, the ICER suffers from the problem that its moments may not be defined, due to the

---

[2] Note that implicit in this assumption of the variance of the cost and effect differences in Equation 3 is that the data where generated from independent groups in a randomised trial. However, for non-randomised trial designs, such as, for example, a before and after study, the independence assumption may not be justified and the above expressions should incorporate a covariance term.

non-negligible probability that the denominator of the ratio could take a zero value.

In contrast to a ratio, where the variance may not be defined, the variance of net-benefits is simply a linear combination of two asymptotically normal variables and can therefore be defined as

$$\text{var}(N\hat{B}) = R_c^2 \text{ var}(\Delta\overline{E}) + \text{var}(\Delta\overline{C}) - 2R_c \text{ cov}(\Delta\overline{E}, \Delta\overline{C}).$$

The advantage of the net-benefits approach is that the $(1-\alpha)\%$ confidence interval for net-benefits can be easily determined in the standard fashion, as $N\hat{B} \pm z_{\alpha/2}\sqrt{\sigma_{NB}^2}$ where $N\hat{B}$, is the estimated net-benefit measure, with variance σ_{NB}^2, and is the critical value from the standard normal distribution.

The assumption of net-benefits following a normal distribution is convenient when it comes to apply Bayesian methods. Bayes theorem can be written as the relationship between the posterior distribution of θ (the vector of parameters of interest) $h(\theta \mid \mathbf{x})$, the prior distribution function $g(\theta)$; and the observed data $\mathbf{X}$ with a density function $f(\mathbf{x} \mid \theta)$ such that

$$h(\theta \mid \mathbf{x}) = \frac{f(\mathbf{x} \mid \theta) g(\theta)}{\int f(\mathbf{x} \mid \theta) g(\theta) d\theta}.$$

Suppose that the normal likelihood is chosen to describe the data density function such that $X \sim N(\mu, \sigma^2)$ and the prior distribution of θ is $N(\tau, \omega^2)$. Putting these distributions into the expression for Bayes theorem given above yields the result that the posterior distribution for θ is given as

$$h(\theta \mid \mathbf{x}) = N\left(\frac{\omega^2\mu + \sigma^2\tau}{\omega^2 + \sigma^2}, \frac{\varpi^2\sigma^2}{\omega^2 + \sigma^2}\right).$$

Defining $W = \omega^2 / (\omega^2 + \sigma^2)$ simplifies the above expression to

$$h(\theta \mid \mathbf{x}) = N\left(W\mu + (1-W)\tau, W\sigma^2\right).$$

If the variance of the prior distribution is very large, then the weight given to the prior mean and variance will be very small. A noninformative prior is equivalent to a very large variance for the prior. Since as $\omega^2 \to \infty$, $W \to 1$, it is clear that a noninformative prior leads to a posterior distribution that is dominated by the data such that $h(\theta \mid \mathbf{x}) = N(\mu, \sigma^2)$.

Reference List

Briggs, A.H. and Fenn, P. (1998) Confidence intervals or surfaces? Uncertainty on the cost-effectiveness plane. *Health Economics* **7**(8) (in press).

Briggs, A.H., Wonderling, D.E. and Mooney, C.Z. (1997) Pulling cost-effectiveness analysis up by its bootstraps: a non-parametric approach to confidence interval estimation. *Health Economics* **6**, 327-340.

Brophy, J.M. and Joseph, L. (1995) Placing trials in context using Bayesian analysis. GUSTO revisited by Reverend Bayes. *JAMA* **273**, 871-875.

Carlin, R.P. and Louis, A.T. (1996) *Bayes and empirical Bayes methods for data analysis.* London: Chapman & Hall.

Chaudhary, M.A. and Stearns, S.C. (1996) Estimating confidence intervals for cost-effectiveness ratios: An example from a randomized trial. *Statistics in Medicine* **15**, 1447-1458.

Claxton, K. and Posnett, J. (1996) An economic approach to clinical trial design and research priority-setting. *Health Economics* **5**, 513-524.

Kass, R.E. and Greenhouse, J.B. (1989) Comments on 'Investigating therapies of potentially great benefit: ECMO' (by J.H.Ware). *Statist.Sci.* **4**, 310-317.

O'Brien, B.J., Drummond, M.F., Labelle, R.J. and Willan, A. (1994) In search of power and significance: issues in the design and analysis of

stochastic cost-effectiveness studies in health care. *Medical Care* **32**, 150-163.

Obenchain, R.L., Melfi, C.A., Croghan, T.W. and Buesching, D.P. (1997) Bootstrap analyses of cost effectiveness in antidepressant pharmacotherapy. *PharmacoEconomics* **11**, 464-472.

Phelps, C.E. and Mushlin, A.I. (1991) On the (near) equivalence of cost-effectiveness and cost-benefit analyses. *Int.J Technol.Assess.Health Care* **7**, 12-21.

Polsky, D., Glick, H.A., Willke, R. and Schulman, K. (1997) Confidence intervals for cost-effectiveness ratios: a comparison of four methods. *Health Economics* **6**, 243-252.

Raikou, K., Gray, A., Briggs, A., Stevens, R., Cull, C., McGuire, A., Fenn, P., Stratton, I., Holman, R. and Turner, R. for the UK Prospective Diabetes Study Group (1998) Cost effectiveness analysis of improved blood pressure control in hypertensive patients with type 2 diabetes: UKPDS 40. *British Medical Journal* **317**, 720-726.

Rice, J.A. (1995) *Mathematical statistics and data analysis*, 2nd Edition. Belmont: Duxbury Press.

Spielgelhalter, D.J., Freedman, L.S. and Parmar, M.K.B. (1994) Bayesian approaches to randomised trials. *Journal of the Royal Statistical Society, Series B* **157**, 357-416.

Stinnett, A.A. and Mullahy, J. (1998) Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. *Medical Decision Making.* **18**, S65-S80.

Tambour, M., Zethraeus, N. and Johannesson, M. (1998) A note of confidence intervals in cost-effectiveness analysis. *International Journal of Technology Assessment in Health Care* 14(3): 467-471.

The GUSTO investigators (1993) An international randomized trial comparing four thrombolytic strategies for acute myocardial infarction. *New England Journal of Medicine* **329**, 673-682.

van Hout, B.A., Al, M.J., Gordon, G.S. and Rutten, F.F. (1994) Costs, effects and C/E-ratios alongside a clinical trial. *Health Economics* **3**, 309-319.

Wakker, P. and Klaassen, M. (1995) Confidence intervals for cost-effectiveness ratios. *Health Economics* **4**, 373-382.

Willan, A.R. and O'Brien, B.J. (1996) Confidence intervals for cost-effectiveness ratios: an application of Fieller's theorem. *Health Economics* **5**, 297-305.

Table 1

Results from a (hypothetical) randomised controlled trial of a new therapy (treatment group) compared to an existing therapy (control group).

	Mean	SE	low 95%	upp 95%	corr
Control Group					
Effect	8.89	0.26	8.39	9.39	
Cost	£25,080	£ 1,058	£23,006	£27,155	0.12
Treatment Group					
Effect	9.62	0.23	9.18	10.07	
Cost	£28,645	£ 1,421	£25,859	£31,430	-0.03
Difference					
Effect	0.74	0.34	0.07	1.41	
Cost	£ 3,564	£ 1,772	£ 91	£ 7,037	0.03
ICER	£ 4,836				
Net Benefit ($R_c = £10{,}000$)	£ 3,806	£ 3,803	-£ 3,648	£11,260	

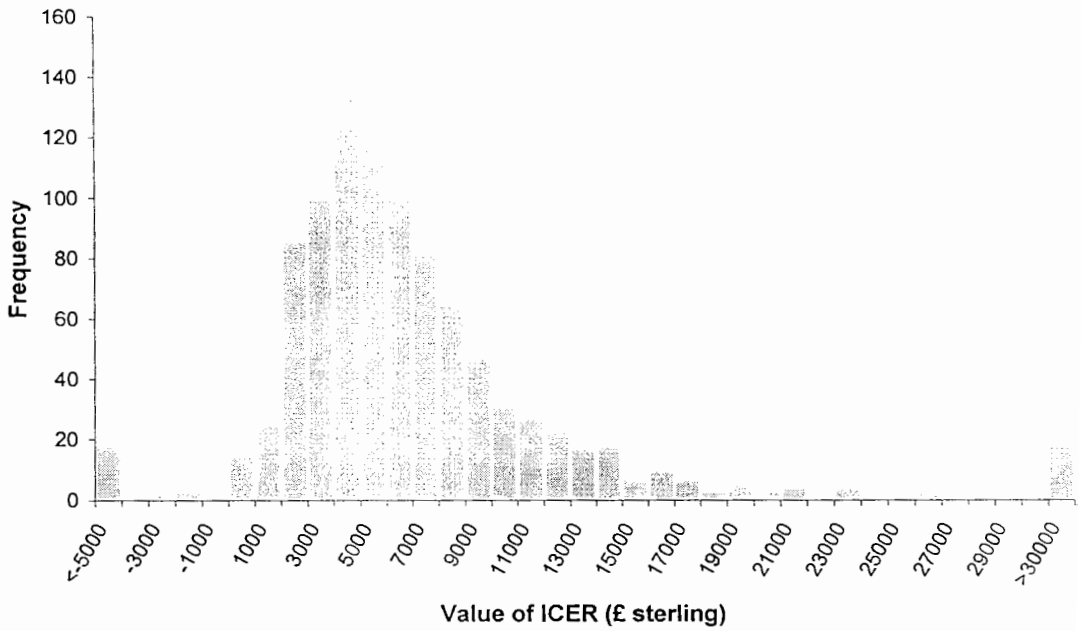
Table 2

Prior information from a previous study

	Mean	SE	low 95%	upp 95%	corr
Control Group					
Effect	8.68	0.52	7.65	9.71	
Cost	£ 27,349	£ 1,372	£ 24,661	£ 30,037	-0.02
Treatment Group					
Effect	10.06	0.50	9.09	11.03	
Cost	£ 28,293	£ 1,494	£ 25,365	£ 31,221	0.12
Difference					
Effect	1.38	0.72	-0.04	2.80	
Cost	£ 944	£ 2,028	-£ 3,031	£ 4,916	0.05
ICER	£ 684				
Net Benefit (R_c = £10,000)	£ 12,856	£ 7,410	-£ 1,667	£ 27,380	

Figure 1

Bootstrap estimate of the sampling distribution of the ICER for the data presented in Table 1



(Note that the highest and lowest values are shown as a single bar in order to present the histogram more clearly.)

Figure 2

Bootstrap replications of cost and effect differences on the cost-effectiveness plane

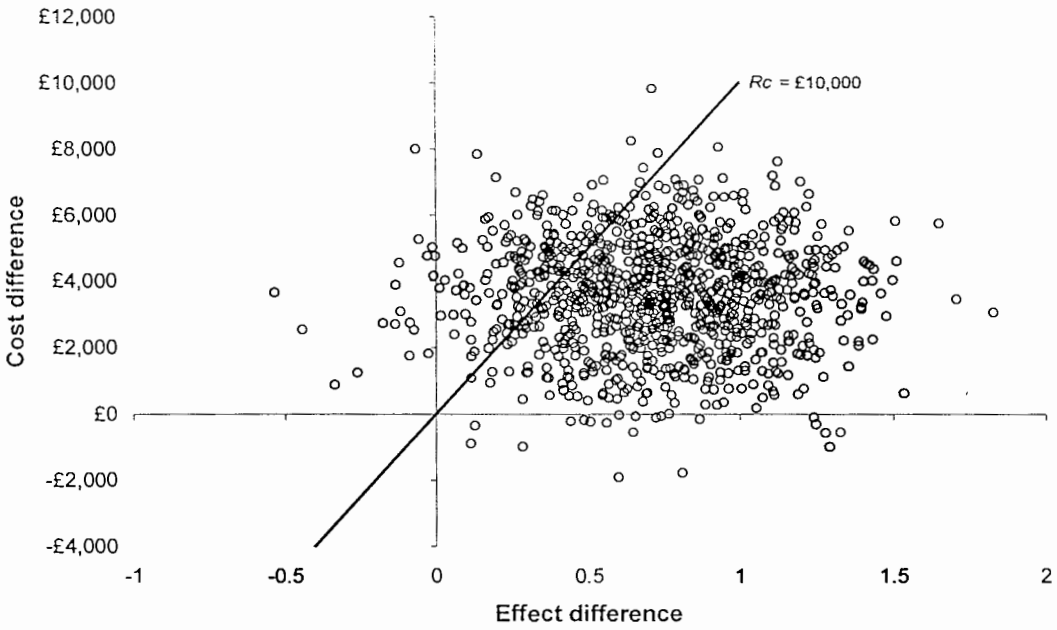
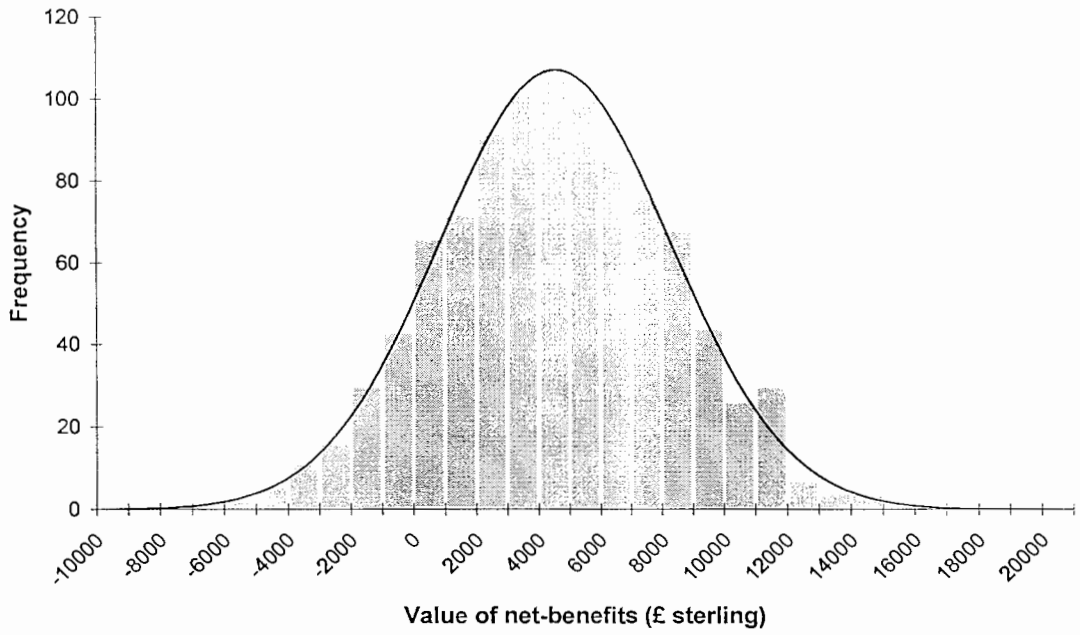


Figure 3

Bootstrap estimate of the sampling distribution of the net-benefits statistic



(Note that a normal distribution with the same mean and variance as the bootstrapped net-benefit data has been overlaid.)

Figure 4

Cost-effectiveness acceptability curve for the data from Table 1

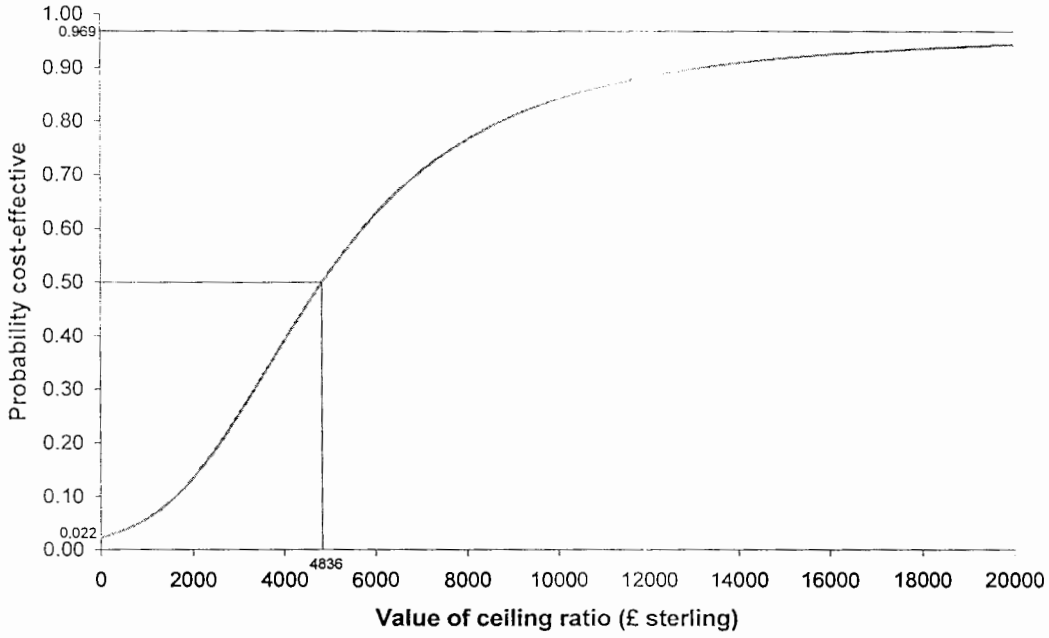


Figure 5

Bayesian approach to cost-effectiveness analysis. Two posterior distributions for net benefits based on an empirical prior distribution (shown) and an uninformative prior.

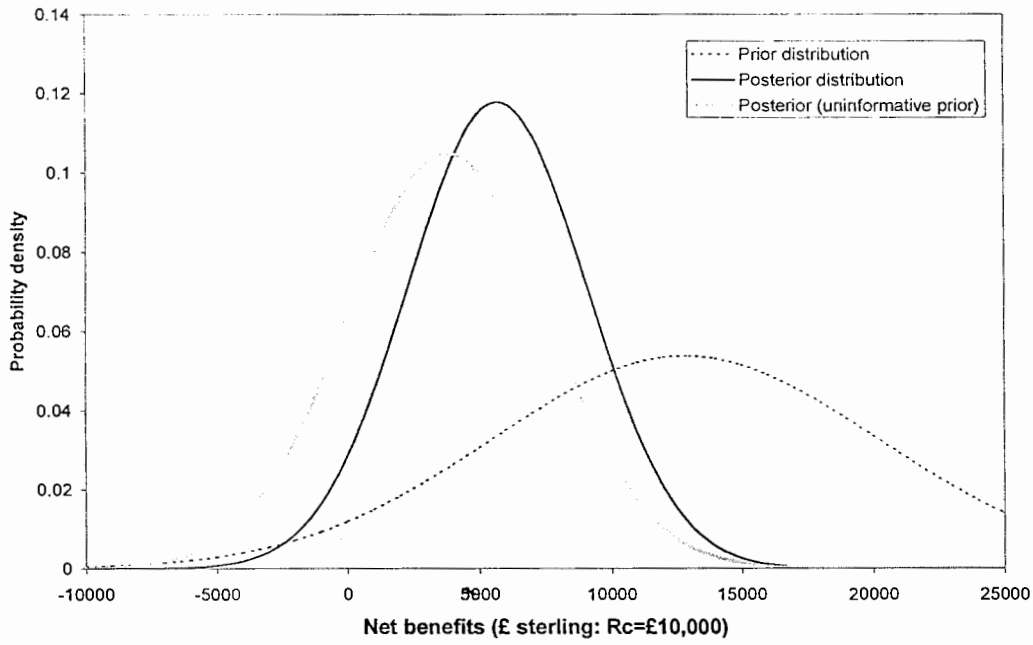


Figure 6

Bayesian approach to cost-effectiveness analysis. Two posterior cost-effectiveness acceptability curves based on an empirical prior distribution (shown) and an uninformative prior.