A BAYESIAN APPROACH TO STOCHASTIC COST-EFFECTIVENESS ANALYSIS

An Illustration and Application to Blood Pressure Control in Type 2 Diabetes

Andrew H. Briggs
University of Oxford

Abstract

The aim of this paper is to discuss the use of Bayesian methods in cost-effectiveness analysis (CEA) and the common ground between Bayesian and traditional frequentist approaches. A further aim is to explore the use of the net benefit statistic and its advantages over the incremental cost-effectiveness ratio (ICER) statistic. In particular, the use of cost-effectiveness acceptability curves is examined as a device for presenting the implications of uncertainty in a CEA to decision makers. Although it is argued that the interpretation of such curves as the probability that an intervention is cost-effective given the data requires a Bayesian approach, this should generate no misgivings for the frequentist. Furthermore, cost-effectiveness acceptability curves estimated using the net benefit statistic are exactly equivalent to those estimated from an appropriate analysis of ICERs on the cost-effectiveness plane. The principles examined in this paper are illustrated by application to the cost-effectiveness of blood pressure control in the U.K. Prospective Diabetes Study (UKPDS 40). Due to a lack of good-quality prior information on the cost and effectiveness of blood pressure control in diabetes, a Bayesian analysis assuming an uninformative prior is argued to be most appropriate. This generates exactly the same cost-effectiveness results as a standard frequentist analysis.

Economic evaluation of healthcare interventions is fundamentally concerned with decision making. In the context of a budget-constrained system of health care, the effectiveness of a healthcare 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 healthcare 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 summarize 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

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 U.K. MRC/South-East Region Training Fellowship in Health Economics and Medical Statistics. Permission to use the U.K. Prospective Diabetes Study results for the hypertension study (UKPDS 40) is gratefully acknowledged.

good value for money. If, however, the ICER is greater than R_c, the intervention is not cost-effective and resources should be prioritized to other more worthwhile interventions.

Of course, estimates of the cost-effectiveness of healthcare 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 the presentation of uncertainty in the results of cost-effectiveness analysis (CEA) more generally (2;4;6;17;19;21;26;27;29).

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 (5;13;14).

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 statistical information available from the study sample. A second Bayesian approach 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 emphasize 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.

In addition to discussing the common ground between Bayesian and frequentist approaches to CEA, a further aim of the paper is to explore the use of the newly developed net benefit statistic (26;27) and its advantages over the traditional ICER. The next section introduces the standard frequentist approach to handling uncertainty in economic evaluation and outlines the 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 CEA. The net benefit statistic is employed in preference to the ICER due to its desirable statistical properties. However, the overall equivalence of the net benefit 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. Initially, illustrative data are employed to elucidate the methods, and this is then followed by a practical application of the Bayesian approach to the cost-effectiveness of tight blood pressure control in hypertensive type 2 diabetic patients in the U.K. Prospective Diabetes Study (UKPDS 40) (22). Since the focus of the paper is on the intuitive appeal of the Bayesian approach rather than the technical aspects, statistical formulae are kept to a minimum.

HANDLING UNCERTAINTY IN COST-EFFECTIVENESS ANALYSIS

In an economic evaluation undertaken alongside a clinical trial, on the basis of data collected from two groups of patients receiving alternative therapies, the ICER can be estimated by

Table 1. Results from a (Hypothetical) Randomized Controlled Trial of a New Therapy (Treatment Group) Compared to an Existing Therapy (Control Group)

	Mean	SE	Lower 95%	Upper 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	

$$\hat{R} = \frac{\bar{C}_T - \bar{C}_C}{\bar{E}_T - \bar{E}_C} = \frac{\Delta\bar{C}}{\Delta\bar{E}} \tag{1}$$

where $\bar{C}_T$ and $\bar{C}_C$ are the mean costs in the treatment and control arms of the trial, respectively, and $\bar{E}_T$ and $\bar{E}_C$ are the mean effects.[1] 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 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. 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 nonparametric bootstrap method (4;6;17) allows the uncertainty in cost-effectiveness to be visualized and emphasizes 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 underlying the summary statistics presented in 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 this case does not follow a well-behaved distributional form. (Indeed, the histogram of Figure 1

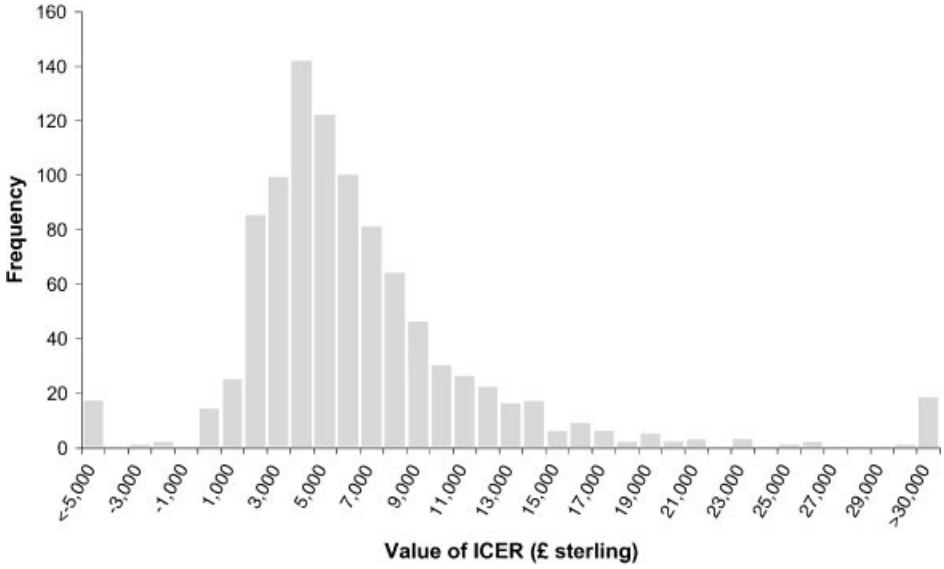

Figure 1. Bootstrap estimate of the sampling distribution of the ICER for the data presented in Table 1. (Note that the highest and lowest values are shown as a single bar in order to present the histogram more clearly.)

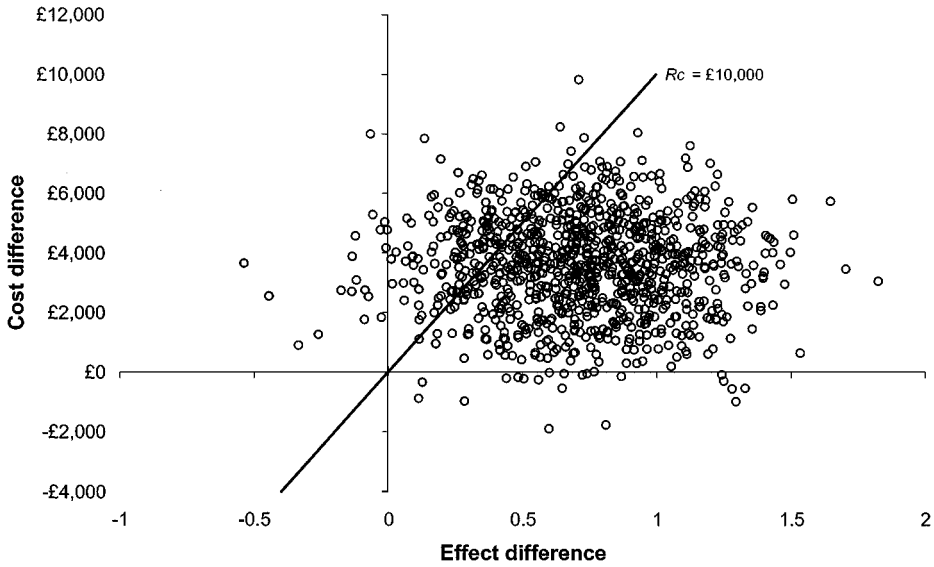

Figure 2. Bootstrap replications of cost and effect differences on the cost-effectiveness plane.

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 northeast quadrant of the plane where the ICER is positive, an important minority of cost/effect pairs fall in the southeast and northwest 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.[2] 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 (2).

Recently, a net benefit approach has been suggested that provides a solution to the problem, providing that the ceiling ratio appropriate for decision making, R_c, is known (26;27). This approach was also employed in an earlier paper by Claxton and Posnett (7) that adopted a Bayesian approach to optimal trial design. 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 (7;27) or the effect scale (26). This is the same formulation adopted by Phelps and Mushlin (20) to argue for the (near) equivalence of cost-benefit and cost-effectiveness analyses, although they did not address the specific issue of uncertainty. Since the net benefit statistics are exactly equivalent, the choice of scale is a matter of preference. For this paper, the net benefit on the cost scale[3] is used and defined as:

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

Note that no problems of interpretation for the net benefit statistic arise: negative effect differences with positive cost differences give a negative net benefit, while positive effect differences and negative cost differences give a positive net benefit. Furthermore, in contrast to the ICER, which does not have a mathematically tractable variance, the variance for net benefit is simply defined in the standard way as:

$$\text{var}(NB) = R_c^2 \cdot \text{var}(\Delta\bar{E}) + \text{var}(\Delta\bar{C}) - 2 \cdot R_c \cdot \text{cov}(\Delta\bar{E}, \Delta\bar{C}).$$

Therefore, the standard statistical approach would be to estimate the confidence interval for net benefit and to see whether that interval excludes zero. In terms of Figure 2, employing $R_c = £10{,}000$ will lead to positive net-benefit for bootstrap replications to the right of the line representing the decision rule and negative net-benefit to the left. A further advantage of the net benefit statistic is its well-behaved statistical properties. Figure 3 shows 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 remain. First, what are the implications for decision making when the net benefit statistic is nonsignificant? For example, Table 1 gives the bootstrap estimate of the 95% confidence interval for the net benefit as $-£3{,}648$ to £11,260, indicating that the intervention is not significantly cost-effective at the 5% level. Secondly, the net benefit statistics rely on a predefined value for R_c, whereas in fact the value of R_c is unknown.

Both these problems can be addressed simultaneously by plotting the cost-effectiveness acceptability curve (28). In terms of net benefit, the curve can be estimated by plotting the

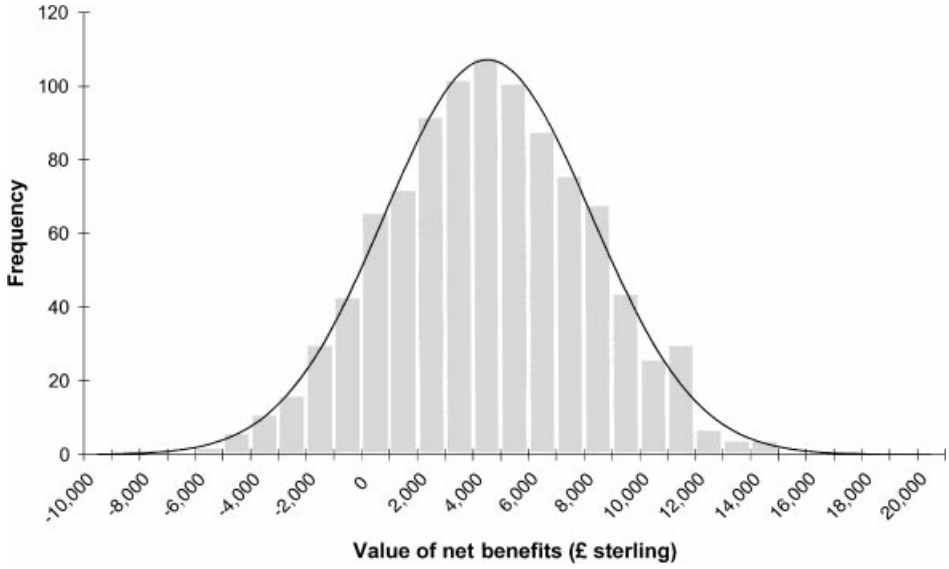

Figure 3. Bootstrap estimate of the sampling distribution of the net benefit statistic. (Note that a normal distribution with the same mean and variance as the bootstrapped net benefit data has been overlaid.)

(one-sided) confidence level at which the net benefit becomes significant as a function of R_c. Although acceptability curves do not solve the problem of the unknown R_c, by conditioning on its unknown value acceptability curves do summarize uncertainty in a way that is directly relevant to the study question of whether the intervention under evaluation is cost-effective.

The acceptability curve was initially described in relation to cost-effectiveness ratios and the cost-effectiveness plane (28), and can be equivalently thought of as considering the proportion of bootstrap replications falling to the right side of the line representing R_c in Figure 2 as the line is rotated counterclockwise from the horizontal ($R_c = 0$) through to the vertical ($R_c = \infty$). Note that since both this approach and the net benefit approach will identify the same bootstrap replications as cost-effective given the value of R_c, the resulting (nonparametric) acceptability curves are (exactly) eqivalent. 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 estimate of the ICER (£4,836) corresponds to the 0.50 acceptability point[4] 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.

A frequentist interpretation of cost-effectiveness acceptability curves is possible by considering that, conditional on the value of R_c, the acceptability curve indicates the (one-sided) p values on the net benefit statistic as described above. However, the natural way to interpret these curves (as illustrated by the labeling of the vertical axis in Figure 4) is as the probability that the intervention is cost-effective, hence providing a direct answer to the study question. Indeed, this is the way cost-effectiveness acceptability curves have been presented in the literature to date (2;22;28), indicating that in the context of (typically) frequentist trials, p values are being interpreted as probabilities. This may reflect that researchers want to be able to make probability statements about the null hypothesis being false (1). Such probability statements are only possible using a Bayesian view of the probability of the hypothesis given the data.

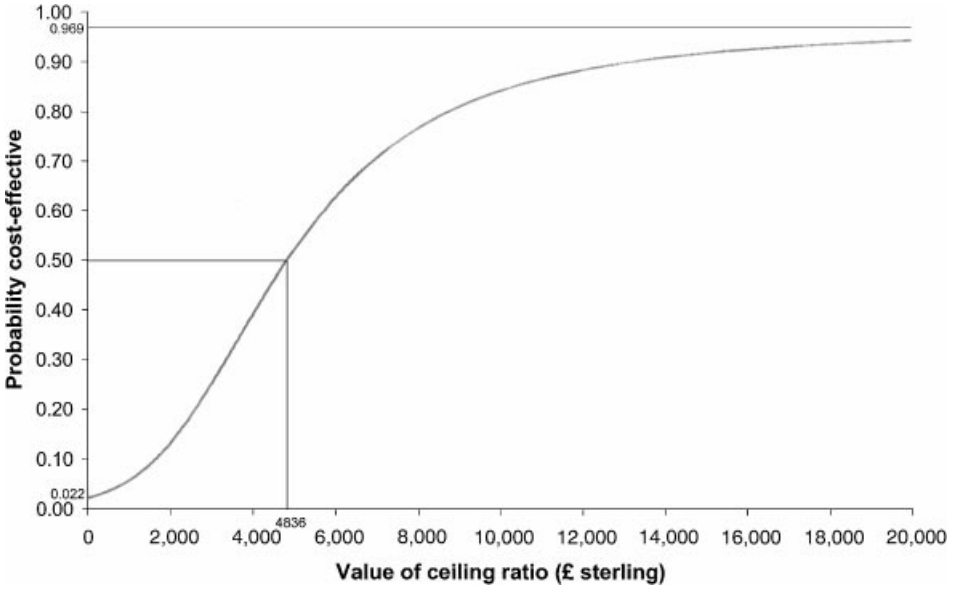

Figure 4. Cost-effectiveness acceptability curve for the data from Table 1.

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 than the traditional long-run frequency based methods, such that frequentists considered Bayesian methods as subjective and highly dependent on prior beliefs, while their own methods were considered 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 (5). The empirical Bayes methods and Bayesian analysis based on uninformative prior distributions are not subjective and have much to offer the frequentist analyst. In this section the standard Bayesian approach is set out in general terms and the updating procedure is shown for normally distributed variables, and a Bayesian analysis of the U.K. PDS trial is illustrated.

Bayes Theorem and the Normal Distribution

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:

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

It is this equation that gives rise to the common expression of the Bayesian posterior distribution as proportional to the prior times the likelihood.

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 (23), which make the calculation of the posterior distribution more straightforward. Although the advent of powerful desktop computers and appropriate software based on Markov chain Monte Carlo methods (such as WinBUGS, developed by the MRC Biostatistics Unit at Cambridge University, U.K., and available for download from their website: www.mrc-bsu.cam.ac.uk/bugs/) has meant that this is no longer strictly necessary, there is added convenience when the calculations can be undertaken in closed analytical form. In particular, the normal distribution is self-conjugate such that normal data and normal prior lead to a normal posterior distribution.

Suppose that a normal likelihood adequately describes the data likelihood function such that $X \sim N(\mu, \sigma^2)$ and the prior distribution of θ can also be expressed in terms of a normal distribution, such as $N(\tau, \omega^2)$. 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:

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

Thus, the mean and variance of the posterior distribution can be seen as a weighted average of the means and variances of the prior and likelihood functions. If the variance of the prior distribution is very large compared with that of the data likelihood, 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 likelihood such that $h(\theta \mid \mathbf{x}) = N(\mu, \sigma^2)$. Thus, in the case of normally distributed likelihood and priors, there will be equivalence between the frequentist approach and the Bayesian posterior arising from the use of a noninformative prior.

For CEA, the net benefit statistic will be much more convenient to handle in a Bayesian analysis than would the ICER statistic, since it is clear from the example given above that the net benefit statistic is well behaved (Figure 3), closely approximating a normal distribution, when the ICER statistic has a highly erratic distribution caused by non-negligible probability of the denominator of the ratio in the neighborhood of zero.

A Bayesian Approach to the Cost-effectiveness of Tight Blood Pressure Control in Diabetic Patients

To illustrate a Bayesian approach to CEA, data from the recently completed UKPDS 40 are employed. These data relate to a CEA designed to assess the efficiency of tight blood pressure control compared with less tight control in hypertensive patients with type 2 diabetes (22). Hypertension in subjects with type 2 diabetes is a risk factor for macrovascular complications. Although improved blood pressure control has been shown to reduce myocardial infarction and stroke in a diabetic subgroup of elderly patients with type 2 diabetes (24), no information on younger patients or the effect on complications of diabetes was available. The authors of the original CEA noted that although cost-effectiveness of antihypertensive programs based on education and drugs has been reported for a number of populations:

> . . .these analyses have mainly been based on models and lack information on effectiveness and use of resources from long term trials, and none has considered hypertensive patients with type II diabetes. (22, p. 720)

Table 2. Summary Statistics for the Cost-effectiveness of Tight Blood Pressure Control Compared to Less Tight Control in the U.K. Prospective Diabetes Study

	Mean	SE	Lower 95%	Upper 95%
Control group				
Effect	10.30	0.17	9.97	10.64
Cost	£6,145	£434	£5,294	£6,996
Treatment group				
Effect	10.63	0.12	10.41	10.96
Cost	£6,381	£309	£5,775	£6,987
Difference				
Effect	0.33	0.21	−0.08	0.73
Cost	£236	£533	−£808	£1,280
ICER	£720	N/A	N/A	N/A
Net benefit ($R_c = £10{,}000$)	£3,041	£2,135	−£1,143	£7,226

The authors then report the results of their study in the form of incremental costs, incremental effects, and employing the cost-effectiveness acceptability curve approach. These results, assuming standard practice costs and discount rates of 6% for both costs and life-years, are reproduced in Table 2. Also included in Table 2 is an assessment of the net benefit of tight blood pressure control assuming a ceiling ratio for decision making of £10,000 per year of life gained. It is clear from Table 2 that although tight blood pressure control looks both more effective (in terms of life-years gained) and more costly, neither of these differences are statistically significant at conventional levels. Furthermore, the net benefit statistic shows a positive net benefit at $R_C = £10{,}000$ but is also nonsignificant. A naïve economic analysis might conclude that there was nothing to choose between these two treatments; however, this would ignore the wealth of information built up in the UKPDS 40 over an 11-year median follow-up time that the effectiveness of treatment is approaching the standard level of statistical significance. Indeed, the clinical endpoint of time to a diabetes-related event was shown to have a statistically significant difference between the two treatments. Instead the authors presented the results of their analysis in terms of a cost-effectiveness acceptability curve showing, for example, that for $R_C = £10{,}000$, there is a 92% probability that the intervention is cost-effective.

Of course, such an interpretation is only possible using a Bayesian view of probability. The quote from the original paper above indicates that the authors of the original study had considered that there was little information on the cost-effectiveness of tight blood pressure control for patients with diabetes prior to their reported results. However, a search of a database of cost-effectiveness analyses reporting cost per QALY and cost per life-year results (3) for economic analyses of hypertension control conducted alongside a clinical trial identified an economic analysis of the Swedish Trial of Old Patients with Hypertension (the STOP Hypertension study) (8;15). Clearly, the population studied (elderly Swedish patients without diabetes) differed from that of hypertensive patients in the UKPDS 40, but in the absence of other information, this trial might be seen as providing at least some information of the likely costs and effects that might be observed in the UKPDS population. Taking the average estimates of the life-years gained and costs reported in the economic analysis of the STOP study suggests that an estimated 0.16 life-year might be gained from treatment at an additional cost of approximately £1,400 (converted from Swedish crowns and inflated to 1996 prices). Unfortunately, there was little information on sampling variation given in the economic analysis of the STOP study, and standard errors on cost

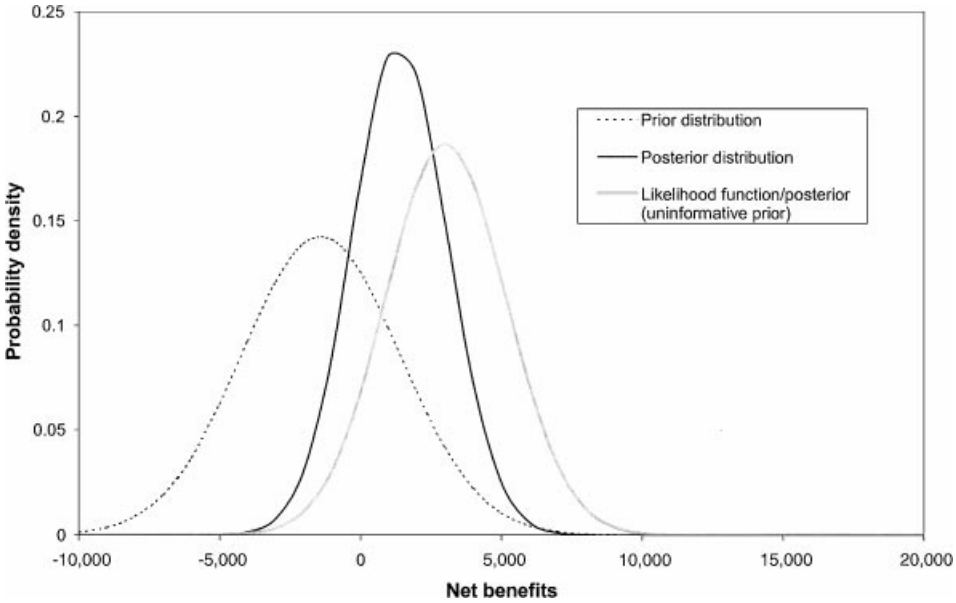

Figure 5. Bayesian approach to cost-effectiveness analysis of tight blood pressure control for hypertensive diabetic patients. Two posterior distributions for net benefit are employed based on a prior distribution from previous studies (shown) and an uninformative prior.

and effect differences were not given. Therefore, the standard errors were set arbitrarily to give a coefficient of variation on cost and effect equal to 2 in order to be conservative and to reflect the fact that the different methods employed in the economic analysis, the difference between the patient populations, and the difference between the two countries' healthcare systems will all increase the level of uncertainty associated with these prior estimates.

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 2. 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—and exactly corresponds with the frequentist notion of likelihood for net benefit for the reasons outlined above. 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 Bayesian approach weights the prior information in relation to its variance compared with the observed data.

Having estimated the prior and posterior distributions for net benefit using Bayesian methods, the probability of the intervention being cost-effective 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. Just as the Bayesian posterior distribution of net benefit corresponds to the frequentist likelihood, so the cost-effectiveness acceptability curve plotted under the assumption of a noninformative prior will correspond exactly to a cost-effectiveness acceptability curve calculated by frequentist methods (involving the less natural interpretation of the curve based on p values).

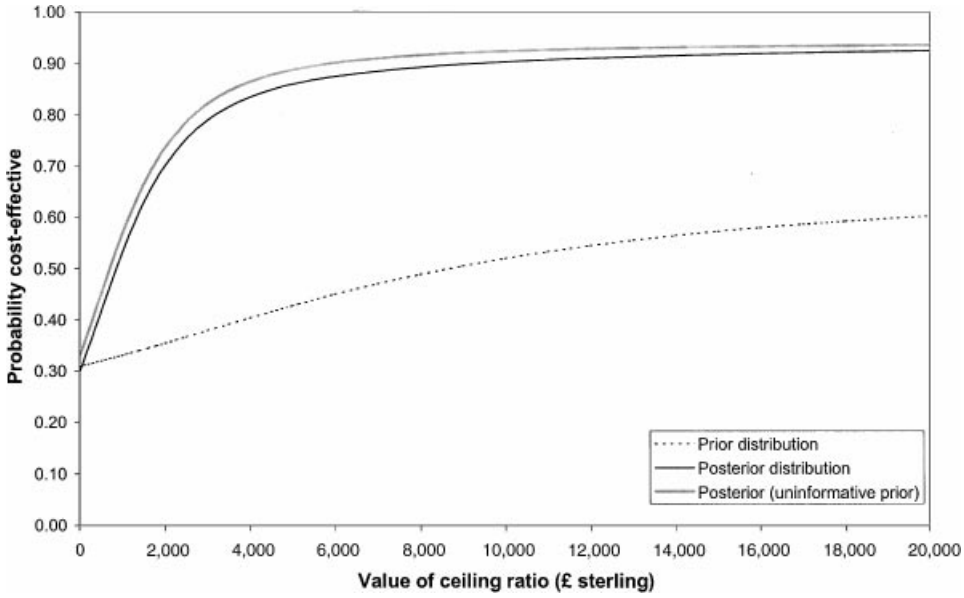

Figure 6. Bayesian approach to cost-effectiveness analysis of tight blood pressure control for hypertensive diabetic patients. Two posterior cost-effectiveness acceptability curves are generated based on a prior distribution from previous studies (shown) and an uninformative prior.

DISCUSSION

In undertaking an economic evaluation of a healthcare intervention, the study question is likely to be: is the intervention cost-effective? Where the evaluation is carried out by means of an experiment with patient-level data generated on costs and effects, it will be natural to consider uncertainty due to sampling variation in a statistical manner. The Bayesian approach offers an important and powerful set of methods to the economic analyst seeking to advise policy decisions because it allows the study question to be addressed directly. Furthermore, 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 through the use of empirical or noninformative priors. In particular, the Bayesian approach allows a more natural way of interpreting cost-effectiveness acceptability curves.

Bayesian methods also provide a natural iterative approach to meta-analysis where the evidence from new trials is simply added as it becomes available (10;11;14), compared with conducting experiments in apparent isolation and then pooling the results, which is the frequentist approach to meta-analysis. Some Bayesians would argue that such an approach is not in fact Bayesian at all since no subjective beliefs are employed. For example, Spiegelhalter and colleagues (25) argue that, while previous results should form the basis of prior distributions, those results should not specify the distribution completely. Referencing Kass and Greenhouse (16), 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.

While a common perception of Bayesian methodology is that it is inherently subjective, some authors have gone to some lengths to emphasize the robustness and objectivity of the application of Bayesian methods (5;14). Furthermore, most Bayesians would be uncomfortable in a situation where the choice of different potential prior distributions has important effects on the posterior distribution (i.e., where the data do not dominate the prior), hence the emphasis in many texts on sensitivity analysis and Bayesian "robustness"

(5;12). Moreover, while subjective priors have a role in personal decision making, it is less clear whether decision makers should be employing subjective beliefs in decisions taken on behalf of society. The recent interest in medical practice having an evidence base could be seen as a way of improving the objectivity of decision making and reducing the influence of personal belief concerning the effectiveness of treatments. What is needed is a "policing" of priors such that if a large number of clinicians disagree as to their prior beliefs concerning the likely effectiveness of an intervention, then this relative uncertainty is included in any Bayesian analysis through the use of a "societal" prior that reflects the underlying uncertainty among experts. Such a composite prior would by definition be relatively noninformative where experts were in disagreement, and therefore any analysis should be largely dominated by the data. By representing diverse individual priors as a vague overall prior, what is potentially a diverse set of posterior distributions from the application of subjective Bayes methods at a personal level becomes a single posterior that incorporates prior uncertainty among experts at the level of society.

At present, and most likely in the immediate 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. A number of commentators have suggested that it may be appropriate for economic analysts to work with error rates (in the frequentist sense) that are higher than those employed in clinical evaluation (9;18). This suggestion indicates the desire of economic analysts to consider the weight of evidence relating to the cost-effectiveness of the intervention under evaluation rather than relying on showing significance at conventional levels. This is most easily achieved through the use of cost-effectiveness acceptability curves that show the weight of evidence for the intervention being cost-effective for all possible values of the ceiling ratio, R_c.

In the original economic evaluation of hypertension control in the UKPDS 40 (22), the authors presented their results in the form of a cost-effectiveness acceptability curve. This paper has demonstrated how prior information on the costs and effects of this intervention could have been incorporated into the analysis. However, this illustration of the Bayesian approach was intended to be illustrative. Given the lack of prior information on the costs and effects of hypertension control for patients with diabetes, the use of a noninformative prior (consistent with the authors' approach and equivalent to a frequentist analysis) was entirely appropriate.

In this paper, the focus has been on estimating the posterior distribution of the parameter of interest (net benefit) 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 minimizes the loss function). Claxton and Posnett (7) 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.

The increasing interest in value for money in the area of health care has resulted in many more clinical trials incorporating an economic analysis as part of their design. As these trials come to be reported, many more analysts are going to face the sorts of problems in representing uncertainty highlighted in this paper. Fortunately, past acrimony between the Bayesian and frequentist schools appears to have been laid aside in favor of a more harmonious coexistence. Two things are clear: a) Bayesian methods allow analysts to address the fundamental question of whether an intervention is cost-effective; and

b) the normative question of how to allocate scarce healthcare resources in a publicly funded healthcare system requires an evidence base such that Bayesian posterior distributions are dominated by the available data rather than by subjective belief.

NOTES

[1] This formulation assumes that patients are homogeneous with respect to risk factors for treatment effect across all patients on the absolute scale. This may be a very strong assumption in many circumstances. A weaker assumption might be that treatment odds ratios are constant, but that absolute risks can vary across patients. Average treatment effect on the absolute scale could then be modeled.

[2] Such a solution would work in this case since there are no positive bootstrap replications in the southwest quadrant of the plane. Note that positive replications in the southwest quadrant follow the opposite ordering from those in the northeast quadrant since high values of the ICER when both costs and effect differences are negative are relatively more favorable to the intervention under evaluation than low ICERs.

[3] The advantages of using the cost scale are that such a formulation corresponds with economists' traditional notion of cost-benefit analysis (20) and that the point estimate of net benefit given in equation 2 is clearly linear in the unknown R_c. However, it has been argued that the use of the health scale to represent net health benefit would have the advantage of emphasizing the *opportunity cost* of resource expenditure directly in terms of health benefits forgone (26).

[4] This correspondence between the ICER and the 0.50 point occurs in this example because it was shown that net benefit is normally (i.e., symmetrically) distributed. In general, this 0.50 point on the acceptability curve will not correspond to the point estimate of the ICER if bootstrapping indicates a skewed distribution for the net benefit.

REFERENCES

1. Bland JM, Altman DG. Bayesians and frequentists. *BMJ*. 1998;317:1151-1160.
2. Briggs AH, Fenn P. Confidence intervals or surfaces? Uncertainty on the cost-effectiveness plane. *Health Econ*. 1998;7:723-740.
3. Briggs AH, Gray AM. Handling uncertainty when performing economic evaluation of health care interventions. *Health Technol Assess*. 1999;3.
4. 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.
5. Carlin RP, Louis AT. *Bayes and empirical Bayes methods for data analysis*. London: Chapman and Hall; 1996.
6. Chaudhary MA, Stearns SC. Estimating confidence intervals for cost-effectiveness ratios: An example from a randomized trial. *Stat Med*. 1996;15:1447-1458.
7. Claxton K, Posnett J. An economic approach to clinical trial design and research priority-setting. *Health Econ*. 1996;5:513-524.
8. Dahlof B, Lindholm LH, Hansson L, et al. Morbidity and mortality in the Swedish Trial in Old Patients with Hypertension (STOP-Hypertension). *Lancet*. 1991;338:1281-1285.
9. Drummond MF, O'Brien B, Stoddart GL, et al. *Methods for the economic evaluation of health care programmes*, 2nd ed. Oxford: Oxford University Press; 1997.
10. Eddy DM, Hasselblad V, Shachter R. A Bayesian method for synthesizing evidence. The Confidence Profile Method. *Int J Technol Assess Health Care*. 1990;6:31-55.
11. Eddy DM, Hasselblad V, Shachter R. An introduction to a Bayesian method for meta-analysis: The confidence profile method. *Med Decis Making*. 1990;10:15-23.
12. Gelman A, Carlin JB, Stern HS, et al. *Bayesian data analysis*. London: Chaapman and Hall; 1995.
13. Goodman SN. Toward evidence-based medical statistics, I: The P value fallacy. *Ann Intern Med*. 1999;130:995-1004.
14. Goodman SN. Toward evidence-based medical statistics, II: The Bayes factor. *Ann Intern Med*. 1999;130:1005-1013.
15. Johannesson M, Dahlof B, Lindholm LH, et al. The cost-effectiveness of treating hypertension in elderly people: An analysis of the Swedish Trial in Old Patients with Hypertension (STOP Hypertension). *J Intern Med*. 1993;234:317-323.

16. Kass RE, Greenhouse JB. Comments on 'Investigating therapies of potentially great benefit: ECMO' (by J.H.Ware). *Stat Sci*. 1989;4:310-317.
17. Obenchain RL, Melfi CA, Croghan TW, Buesching DP. Bootstrap analyses of cost-effectiveness in antidepressant pharmacotherapy. *PharmacoEconomics*. 1997;11:464-472.
18. O'Brien B. Economic evaluation of pharmaceuticals: Frankenstein's monster or vampire of trials? *Med Care*. 1996;34(suppl 12):DS99-DS108.
19. O'Brien BJ, Drummond MF, Labelle RJ, Willan A. In search of power and significance: Issues in the design and analysis of stochastic cost-effectiveness studies in health care. *Med Care*. 1994;32:150-163.
20. Phelps CE, Mushlin AI. On the (near) equivalence of cost-effectiveness and cost-benefit analyses. *Int J Technol Assess Health Care*. 1991;7:12-21.
21. Polsky D, Glick HA, Willke R, Schulman K. Confidence intervals for cost-effectiveness ratios: A comparison of four methods. *Health Econ*. 1997;6:243-252.
22. Raikou M, Gray A, Briggs A, et al. Cost effectiveness analysis of improved blood pressure control in hypertensive patients with type 2 diabetes: UKPDS 40. *BMJ*. 1998;317:720-726.
23. Rice JA. *Mathematical statistics and data analysis*, 2nd ed. Belmont: Duxbury Press; 1995.
24. SHEP Cooperative Research Group. Prevention of stroke by antihypertensive drug treatment in older persons with isolated systolic hypertension. Final results of the Systolic Hypertension in the Elderly Program (SHEP). *JAMA*. 1991;265:3255-3264.
25. Spielgelhalter DJ, Freedman LS, Parmar MKB. Bayesian approaches to randomised trials. *J R Stat Soc B*. 1994;157(part 3):357-416.
26. Stinnett AA, Mullahy J. Net health benefits: A new framework for the analysis of uncertainty in cost-effectiveness analysis. *Med Decis Making*. 1998;18(suppl 2):S65-S80.
27. Tambour M, Zethraeus N, Johannesson M. A note on confidence intervals in cost-effectiveness analysis. *Int J Technol Assess Health Care*. 1998;14:467-471.
28. 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.
29. Wakker P, Klaassen M. Confidence intervals for cost-effectiveness ratios. *Health Econ*. 1995; 4:373-382.