Debate

The probability of cost-effectiveness

Anthony O'Hagan*[1] and John W Stevens[2]

Address: [1]Centre for Bayesian Statistics in Health Economics, University of Sheffield, Sheffield, UK and [2]Clinical Sciences, AstraZeneca R&D Charnwood, Loughborough, UK

E-mail: Anthony O'Hagan* - a.ohagan@sheffield.ac.uk; John W Stevens - John.W.Stevens@astrazeneca.com

*Corresponding author

Abstract

Background: The study of cost-effectiveness comparisons between competing medical interventions has led to a variety of proposals for quantifying cost-effectiveness. The differences between the various approaches can be subtle, and one purpose of this article is to clarify some important distinctions.

Discussion: We discuss alternative measures in the framework of individual, patient-level, incremental net benefits. In particular we examine the probability of cost-effectiveness for an individual, proposed by Willan.

Summary: We argue that this is a useful addition to the range of cost-effectiveness measures, but will be of secondary interest to most decision makers. We also demonstrate that Willan's proposed estimate of this probability is logically flawed.

Background

The study of cost-effectiveness comparisons between competing medical interventions has led to a variety of proposals for quantifying cost-effectiveness. Although the most widely used measure is still the incremental cost-effectiveness ratio (ICER), there is increasing preference for the cost-effectiveness acceptability curve (CEAC). Willan [1] has recently proposed an alternative that he calls the probability of cost-effectiveness.

The differences between these various approaches can be subtle, and further complexity is introduced by some authors preferring a Bayesian formulation over more traditional frequentist analysis. One purpose of this article is to clarify some important issues, concerning (a) the perspectives of different decision makers and (b) the distinction between the true value of an unknown parameter and a statistical inference about that parameter.

Discussion

We first review various approaches to measuring cost-effectiveness, including the ICER, the mean incremental net benefit, and the measure proposed by Willan [1]. We then contrast these measures and argue that Willan's proposal is of only secondary interest to a health care provider. All of the cost-effectiveness measures are in practice unknown parameters that must be estimated from data, and we next consider inference about these measures from both the frequentist and Bayesian approaches. Finally, a fundamental flaw in the estimator proposed by Willan [1] for his probability of cost-effectiveness is exposed.

Measures of cost-effectiveness

We consider two competing treatments, drugs, or other health technologies, which we refer to as Treatment 1 and Treatment 2. Conventionally, Treatment 1 is often the standard treatment whereas Treatment 2 is a new or comparator treatment. In reality there will usually be far more than two competing treatments for any condition, but for

the purpose of this article it is enough to consider, like Willan [1], just two treatments.

A little notation is necessary. Let C_i be the cost associated with an individual patient when given Treatment i, and let E_i be the value of an appropriate effectiveness measure associated with that patient when given Treatment i. Now it is important to recognise variation between patients. One patient will incur different costs and experience different effectiveness from another. Therefore, C_i and E_i are random quantities, which we interpret as the cost and effectiveness under Treatment i for an individual patient *randomly* drawn from the population of all patients under consideration. The probability distributions of these random quantities describes how they vary over the population.

In order to compare cost-effectiveness between the two treatments, we require a way to link costs to effectiveness, and this is done through a decision-maker's *willingness to pay* coefficient K. Formally, the decision-maker is prepared to pay K units of money to obtain one unit of effectiveness. Therefore, the *net benefit* of Treatment i for an individual (random) patient is

$$B_i(K) = K E_i - C_i.$$

This expresses net benefit on the monetary scale by converting the E_i units of effectiveness into $K E_i$ units of money before subtracting the cost C_i. (We could equally express net benefit on the effectiveness scale as $E_i - C_i/K$, but the two approaches are clearly formally equivalent.) The notation also emphasises the dependence of the net benefit on the decision-maker's willingness to pay coefficient K.

Treatment 2 would be clearly more cost-effective than Treatment 1 for an individual (random) patient if $B_2(K) > B_1(K)$. This can be expressed simply in terms of the *individual incremental net benefit* (individual INB)

$$D_B(K) = B_2(K) - B_1(K) = K D_E - D_C,$$

where $D_E = E_2 - E_1$ and $D_C = C_2 - C_1$ are the increments in effectiveness and cost, respectively.

If all patients were the same, and experienced the same costs and effectiveness, then the individual INB would be the same for all patients, and could then be called *the* INB. Then the comparison of treatments would become trivial. The INB would quantify the gain (if positive) or loss (if negative) per patient that would result from switching from Treatment 1 to Treatment 2. Treatment 2 would clearly be more cost-effective than Treatment 1 if, and only if, the INB was positive.

However, patients will vary, and the consequence of this is that individual INB will vary between patients, and there is no single value to represent the comparison between the two treatments. Across the population, there is a probability distribution of individual INB.

The measures of cost-effectiveness that are in widespread use in health economics are based on the *mean* of this distribution. We denote the population *mean incremental net benefit* (mean INB) by $\Delta_B(K)$. The standard notation in probability theory for a mean or *expected* value is $\mathcal{E}$, so the mean incremental net benefit is

$$\Delta_B(K) = \mathcal{E}(D_B(K)) = K \Delta_E - \Delta_C,$$

where $\Delta_E = \mathcal{E}(D_E)$ and $\Delta_C = \mathcal{E}(D_C)$ are the population mean increments in effectiveness and cost. Then Treatment 2 is defined to be more cost-effective than Treatment 1, in terms of the population mean, if $\Delta_B(K) > 0$.

The *incremental cost-effectiveness ratio* (ICER) can be expressed as

$$\rho = \Delta_C / \Delta_E,$$

and we can see that $\Delta_B(K) > 0$, i.e. Treatment 2 is more cost-effective than Treatment 1, if $\rho < K$ and $\Delta_E > 0$, or if $\rho > K$ and $\Delta_E < 0$.

The probability of cost-effectiveness as proposed by Willan [1] is the probability that an individual (random) patient will have a positive individual INB. We can denote this by $\theta(K) = Pr(D_B(K) > 0)$. It can also be seen as the proportion of all patients in the population who have positive individual INBs.

$\Delta_B(K)$ and $\theta(K)$ are just two summary measures of the distribution of net benefit in the population. If the distribution is symmetric about its mean, as shown for instance in Figure 1, then the two measures will be in agreement, in the sense that $\Delta_B(K)$ will be positive if and only if $\theta(K)$ is greater than 0.5.

Thus, the distribution represented by the solid curve in Figure 1 has mean $\Delta_B(K) = 1.3$ and $\theta(K) = 0.903$, so that Treatment 2 is more cost-effective in terms of having a higher mean INB and the proportion of patients who will achieve a higher individual INB under Treatment 2 is 90.3%. Conversely, the distribution represented by the dashed curve has $\Delta_B(K) = -0.7$ and $\theta(K) = 0.242$, so the mean INB under Treatment 2 is now less than under Treatment 1, and only 24.2% of patients will obtain a higher individual INB under Treatment 2.
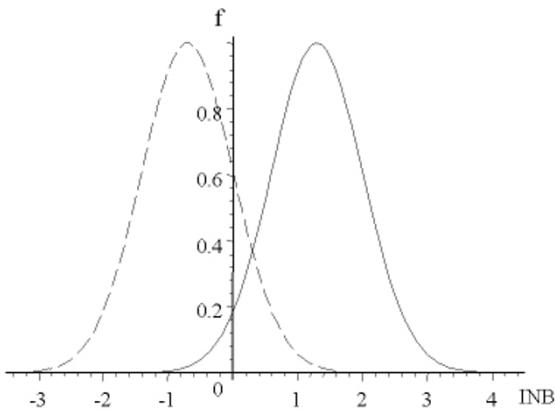
Figure 1
Two symmetric distributions of net benefit.
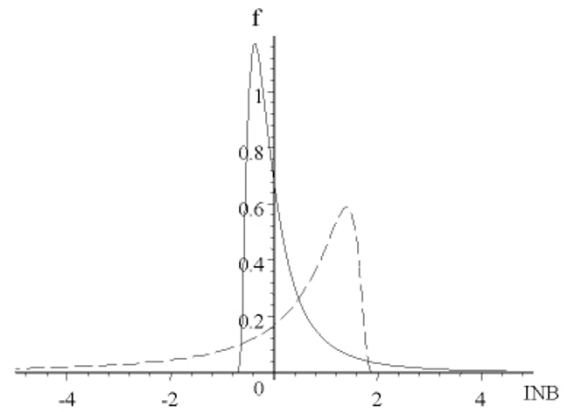
Figure 2
Two skewed distributions of net benefit.

If, however, the distribution is not symmetric, then it is quite possible for the two measures to give apparently contradictory indications of relative cost-effectiveness. Figure 2 shows another two possible distributions. In the distribution shown as a solid line, $\Delta_B(K) = 0.2$ and $\theta(K) = 0.414$, so the mean INB is positive but only 41.4% of patients actually have a higher individual INB under Treatment 2. This is because those 41.4% include an appreciable proportion who obtain large positive individual INBs of 2 or more, whereas although the other 58.6% have negative individual INBs they never experience a value beyond -1. Conversely, in the distribution shown as a dashed line in Figure 2, $\Delta_B(K) = -1$ and $\theta(K) = 0.682$, so that the mean INB is negative but 68.2% of patients have a positive individual INB.

Which measure is best?
It is well-known in health economics that, from the perspective of a health care provider needing to decide which treatment to apply to the population of patients in their care, it is the mean cost and effectiveness over the whole population that matters [2]. This is because the decision is to apply to the whole population. The health care provider will have to pay a cost equal to the total of all the costs for individual patients under the chosen treatment, and when expressed on a per-patient basis this is the population mean cost. For a similar reason, the per-patient mean effectiveness under the chosen treatment measures the benefit that the health care provider obtains for that cost in terms of improved health for the patients in its care. If the health care provider's willingness to pay coefficient is K, then the appropriate measure of relative cost-effectiveness is the mean INB $\Delta_B(K)$, and the correct decision is to fund Treatment 2 if $\Delta_B(K) > 0$ or Treatment 1 if $\Delta_B(K) < 0$ [3].

As discussed in the previous section, this can be expressed in terms of comparing the ICER ρ with K, but that approach is more complex, since the comparison depends on the sign of Δ_E.

From the perspective of a health care provider, then, needing to make a decision between two treatments, the decision rests on mean INB, and in fact only on its *sign*. There is no role for Willan's $\theta(K)$. As we have seen in Figure 2, the wrong decision could be made if it were based on $\theta(K)$.

Willan [1] says, "The use of $\theta(K)$ should be helpful to policy-makers". We agree, in the sense that it does give extra information about the distribution of individual INBs in the population, but as such it is of secondary interest, only. It should not be used as the basis of the actual decision. Nevertheless, we believe that in general an understanding of the distribution of individual INBs in the population is useful ancillary information that may be helpful to a decision-maker in the subsequent implementation of the decision.

The perspective of a health care provider is not necessarily the only one of interest. An individual clinician wishing to decide how to treat an individual patient may be willing to regard that patient as randomly drawn from a large population, and might be interested in $\theta(K)$. However, the situations shown in Figure 2 argue for caution. Consider for instance the dashed curve. The patient is substantially more likely to have a positive individual INB than a negative one, and this may seem to suggest prescribing Treatment 2. There is, however, a risk of a large negative INB, corresponding to the patient having a very much worse outcome with Treatment 2 than with Treatment 1. In our opinion, the mean INB is as relevant to an individ-

ual decision as to the group decision of a health care provider.

Inference about cost-effectiveness

The measures of cost-effectiveness described in the preceding section are all unknown in practice because they depend on the unknown distribution of individual INBs for patients in the population. From the statistical point of view they are unknown parameters. In order to learn about them, we will need to obtain some relevant evidence. This might, for instance, as supposed in Willan [1], consist of observations of actual costs and effectiveness for a sample of patients in a clinical trial.

We then need to construct appropriate methods of statistical inference for parameters of interest, based on the data. There is a substantial literature on this topic. Based on data from a clinical trial, various authors have presented estimators and confidence intervals for the ICER [4–12], and comparable inferences for the mean INB [3,13]. All of these references employ the frequentist approach to statistical inference. Analyses under a Bayesian approach have also been given [14–18]. The fact that the ICER is a ratio, together with the way its interpretation changes as the sign of Δ_E changes, mean that inference about the mean INB is generally much more straightforward [17,18].

Inference about the mean INB is generally presented by means of a *Cost-Effectiveness Acceptability Curve* (CEAC) [16–22]. As introduced by van Hout et al [19], the CEAC plots the probability that mean INB is positive against K. The value of such a graph lies partly in the difficulty of specifying K in practice. Decision-makers are generally reluctant to commit themselves to an explicit willingness to pay, and plotting against K allows them to assess the relative cost-effectiveness of the two treatments over a range of values of K. Strictly, the probability that $\Delta_B(K)$ is positive can only be a Bayesian inference, since only in the Bayesian approach is it possible to assign probability distributions to unknown parameters. The frequentist analogue is to consider the P-value of a significance test of the null hypothesis that $\Delta_B(K) < 0$ [17,23]. It is important to remember, as always, that the interpretation of a P-value for a null hypothesis is much less direct and meaningful than the Bayesian probability that the hypothesis is true.

In its more natural Bayesian form, the CEAC states, for given K, the probability, *based on the available evidence*, that the true value of the unknown parameter $\Delta_B(K)$ is positive. It therefore states, for given K, the probability, *based on the available evidence*, that Treatment 2 is more cost-effective than Treatment 1, from the perspective of a health care provider needing to make a decision between the two treatments. In presenting the CEAC in practice, authors have tended to assert that the CEAC states, for given K, the probability that Treatment 2 is more cost-effective than Treatment 1, omitting to refer to the fact that this probability is based on available evidence, and omitting to state the decision context. Willan [1] objects to this presentation of the CEAC as giving 'the probability of cost-effectiveness'. He writes:

"The interpretation that the acceptability curve is the probability that the intervention is cost-effective is not entirely accurate and could easily be misunderstood by policy makers. Consider the situation in which the observed INB for treatment is very small, but due to a very large sample size the acceptability curve at the value of λ [our K] of interest is 0.99. Attaching the label "the probability that the intervention is cost-effective" to this quantity could mislead policy makers into thinking that treatment is highly beneficial compared to the standard. What, in fact, is high is our confidence that the INB, however small, is not zero."

We agree that to refer to the CEAC as simply 'the probability of cost-effectiveness', or 'the probability that Treatment 2 is more cost-effective than Treatment 1', is potentially misleading if its dependence on the available evidence and on the decision context is not clear. We advocate that the phrase 'based on available evidence' should be used to emphasise the first point, or for a technical audience the Bayesian formulation of 'the posterior probability of cost-effectiveness' would be appropriate. It might be helpful also to emphasise that we are judging cost-effectiveness from the perspective of a health care provider needing to decide between two treatments, although this context has been so pervasively adopted in health economics that we believe it can be taken as understood.

Willan proposes that $\theta(K)$ should more properly be called 'the probability of cost-effectiveness', but to use the phrase for $\theta(K)$ without further qualification would be *equally* misleading to policy makers. To parallel the above quotation from Willan [1], consider the situation in which the mean INB is positive but small, but due to there being very little between-patient variation we find $\theta(K) = 0.99$. Attaching the label 'the probability that the intervention is cost-effective' to this quantity could mislead policy makers into thinking that treatment is highly beneficial compared to the standard. What, in fact, is high is the proportion of patients in the population for whom the individual INB, however small, is positive.

Neither measure asserts the degree to which one treatment is 'highly beneficial' compared to the other. *Both* are concerned only with the sign of INB. The CEAC gives the probability, based on available evidence, that the mean INB is positive, while $\theta(K)$ gives the probability that an individual INB is positive.

Willan [1] further objects to the fact that the CEAC changes as we get more information, because it is a statistical inference. As we get more evidence, our uncertainty about the sign of $\Delta_B(K)$ for a given value of K will decrease until we become certain either that $\Delta_B(K)$ is positive (whereupon the CEAC will tend to 1) or that it is negative (in which case the CEAC will tend to 0). This is entirely natural, and we do not understand this objection.

Willan's $\theta(K)$ is a probability in a different sense, because it is a population parameter, not an inference about a population parameter. Inferences change as we get more data, while the true values of the underlying parameters remain fixed, but unknown. This does not make $\theta(K)$ in any sense a superior kind of probability. It happens that Willan is interested in inference about a parameter that can itself be considered as a probability (although we believe it would be more helpful to call it a proportion, i.e. the proportion of patients in the population with positive individual INBs). To make inference about it, he provides an estimator (although, as we shall see below, that estimator is logically flawed), but he could have considered calculating a P-value for the null hypothesis that $\theta(K) > 0.5$. That would be analogous to the CEAC, and would change with the available data in the same way.

Willan's estimator
Willan proposes an estimator of $\theta(K)$ based on data comprising observed costs and effectiveness values for a sample of n_S patients given the standard, treatment 1, and another sample of n_T patients given the intervention, treatment 2. Now since these data do not include any observations in which the same patient is given both treatments, it is completely impossible to learn the true value of $\theta(K)$, no matter how large n_S and n_T might be.

It is easy to demonstrate this impossibility with a simple example. Suppose that we have enormous samples such that we learn the true distribution in the population of costs and effects for treatment 1 and the true distribution of costs and effects for treatment 2. In particular, we will also learn the true distribution of net benefits $B_i(K)$ for each treatment. Suppose that the value of K is given and that the distribution of the net benefit $B_1(K)$ under treatment 1 is $N(0,1)$ (i.e. normal with mean 0 and variance 1), while the distribution of net benefit $B_2(K)$ under treatment 2 is $N(1,1)$. With all this information we know these distributions, and so we know exactly that $\Delta_B(K) = 1$. We therefore know with certainty that treatment 2 is more cost-effective than treatment 1 for a health care provider with the given value of K.

Even with all this information we do not know $\theta(K)$, because this depends on how correlated $B_1(K)$ and $B_2(K)$ are in the population. At one extreme, they might be perfectly

positively correlated, such that for every individual in the population it is true that $B_2(K) = B_1(K) + 1$. Then $\theta(K) = 1$, because the individual INB is positive for *every* patient. At the other extreme we might have perfect negative correlation, so that for every individual in the population we have $B_2(K) = 1 - B_1(K)$. Then treatment 2 is more cost-effective for all those individuals for whom $B_1(K) < 0.5$. The proportion of such individuals in the population is 69.15%, and so $\theta(K) = 0.6915$.

Willan's estimator effectively assumes that $B_1(K)$ and $B_2(K)$ are independent in the population, and for our example this implies that $D_B(K)$ is distributed as $N(1,2)$, with the result that $\theta(K) = 0.7633$. The assumption is arbitrary and completely unsupported. Indeed one might imagine that in practice there would be quite strong correlations, on the basis that a patient who responds well to one treatment might respond relatively well to the other, and similarly for costs. But we reiterate that there is absolutely no evidence about this correlation in the data which Willan supposes are available. Indeed for most kinds of intervention it is impractical to test two treatments on the same patient, and even when this is possible we must expect the picture to be complicated by cross-over effects.

What Willan [1] actually estimates is the probability that a randomly chosen patient given treatment 1 will obtain a higher net benefit than *another* randomly chosen patient given treatment 2. This is an entirely different measure from $\theta(K)$ and we cannot imagine that it is of fundamental interest to any policy maker.

Summary
1. From the perspective of a health care provider needing to decide which of two treatments to fund, it is the mean cost and mean effectiveness, over the whole population of patients within the provider's remit, that are of primary concern. This leads to the mean INB $\Delta_B(K)$ as the appropriate measure of cost-effectiveness, and to the specific question of whether $\Delta_B(K)$ is positive.

2. Any measure of cost-effectiveness is a property of the population of patients under consideration, and is an unknown parameter. We make statistical inferences about parameters, based on available evidence. The true value of the parameter is fixed, independent of the available evidence, but unknown. Any statistical inference statement about the parameter is liable to change as the evidence changes. The CEAC plots the probability, based on available evidence, that $\Delta_B(K) > 0$, and is the most relevant inference for a health care provider needing to decide between two treatments. Because it is an inference, the CEAC depends on the data.

3. When reporting the CEAC in practice, its dependence on the data should be made clear by referring to it in such phrases as 'the probability of cost-effectiveness based on available evidence' or 'the posterior probability of cost-effectiveness'. It may also be useful to emphasise that cost-effectiveness is being judged from the perspective of a health care provider needing to decide which of two treatments to fund.

4. Willan's probability of cost-effectiveness $\theta(K)$ may be useful to a decision maker in the same way as knowing other aspects of the distribution of individual INBs in the population would be useful, but it will generally be of secondary importance to the sign of $\Delta_B(K)$. Since $\theta(K)$ is a parameter it does not depend on the data, but it is unknown, and any statistical inference about it will depend on the data.

5. $\theta(K)$ should not be referred to simply as 'the probability of cost-effectiveness' either, and we advise calling it, for example, 'the proportion of patients for whom the treatment is cost-effective'. The fact that it is an unknown parameter should be emphasised, by a formulation such as 'based on available evidence, the proportion ... is estimated to be ...'

6. The proposed estimator of $\theta(K)$ given by Willan [1] is flawed. This parameter cannot be estimated consistently from the kind of data considered by Willan. His proposed estimate is in fact a probability concerning two randomly selected future patients, and is of doubtful interest to any decision maker.

Conclusion
In conclusion, therefore, we reiterate the appropriateness of the CEAC as the primary comparator of relative cost-effectiveness between two treatments from the perspective of a health care provider. Willan's 'probability of cost-effectiveness' would be of only secondary value in evidence presented to policy makers, and his proposed estimator of that probability is fatally flawed. However, we agree with Willan that assessments of cost-effectiveness should be more clearly stated, avoiding the unqualified phrase 'the probability of cost-effectiveness'.

Competing interests
None declared.

References
1. Willan AR: **On the probability of cost-effectiveness using data from randomised clinical trials.** *BMC Medical Research Methodology* 2001, **1**:8 [http://www.biomedcentral.com/1471-2288/1/8]
2. Thompson SG, Barber JA: **How should cost data in pragmatic randomised trials be analysed?** *British Medical Journal* 2000, **320**:1197-1200
3. Stinnett AA, Mullahy J: **Net health benefits: A new framework for the analysis of uncertainty in cost-effectiveness analysis.** *Med Decis Making* 1998, **18**:S65-S80
4. Briggs AH, Mooney CZ, Wonderling DE: **Constructing confidence intervals for cost-effectiveness ratios: An evaluation of parametric and non-parametric techniques using monte carlo simulation.** *Statistics in Medicine* 1999, **18**:3245-3262
5. Briggs AH, Wonderling DE, Mooney CZ: **Pulling cost-effectiveness analysis up by its bootstraps: A non-parametric approach to confidence interval estimation.** *Health Economics* 1997, **6**:327-340
6. Chaudhary MA, Steams SC: **Estimating confidence intervals for cost-effectiveness ratios: An example from a randomised trial.** *Statistics in Medicine* 1996, **15**:1447-1458
7. Laska EM, Meisner M, Siegel C: **Statistical inference for cost-effectiveness ratios.** *Health Economics* 1997, **6**:229-242
8. O'Brien BJ, Drummond MF, Labelle RJ, Willan AR: **In search of power and significance: Issues in the design and analysis of stochastic cost-effectiveness studies in health care.** *Medical Care* 1994, **32**:150-163
9. Polsky D, Glick HA, Wilike R, Schulman K: **Confidence intervals for cost-effectiveness ratio: A comparison of four methods.** *Health Economics* 1997, **6**:243-252
10. Tambour M, Zethraeus N: **Bootstrap confidence intervals for cost-effectiveness ratios: Some simulation results.** *Health Economics* 1998, **7**:143-147
11. Wakker P, Klaassen MP: **Confidence intervals for cost-effectiveness ratios.** *Health Economics* 1995, **4**:373-381
12. Willan AR, O'Brien BJ: **Confidence intervals for cost-effectiveness ratios: an application of Fieller's theorem.** *Health Economics* 1996, **5**:297-305
13. Tambour M, Zethraeus N, Johannesson M: **A note on confidence intervals in cost-effectiveness analysis.** *International Journal of Technology Assessment in Health Care* 1998, **14**:467-471
14. Briggs AH: **A Bayesian approach to stochastic cost-effectiveness analysis.** *Health Economics* 1999, **8**:257-261
15. Heitjan DF, Moskowitz AJ, Whang W: **Bayesian estimation of cost-effectiveness ratios from clinical trials.** *Health Economics* 1999, **8**:191-201
16. O'Hagan A, Stevens JW: **A framework for cost-effectiveness analysis from clinical trial data.** *Health Econ* 2001, **10**:302-315
17. O'Hagan A, Stevens JW, Montmartin J: **Inference for the cost-effectiveness acceptability curve and cost-effectiveness ratio.** *PharmacoEconomics* 2000, **17**:339-349
18. Fenwick E, Claxton K, Sculpher M: **Representing uncertainty: The role of cost-effectiveness acceptability curves.** *Health Economics* 2001, **10**:779-787
19. van Hout BA, Al MJ, Gordon GS, Rutten F: **Costs, effects and C/E ratios alongside a clinical trial.** *Health Econ* 1994, **3**:309-319
20. O'Hagan A, Stevens JW, Montmartin J: **Bayesian cost-effectiveness analysis from clinical trial data.** *Statistics in Medicine* 2001, **20**:733-753
21. Raikou M, Gray A, Briggs A, Stevens R, Cull C, McGuire A, Fenn P, Stratton I, Holman R, Turner R, on behalf of the UK Prospective Diabetes Study Group: **Cost effectiveness analysis of improved blood pressure control with type 2 diabetes: UKPDS 40.** *BMJ* 1998, **317**:720-726
22. Sculpher M, Poole L, Cleland J, Drummond M, Armstrong PW, Horowitz JD, Massie BM, Poole-Wilson PA, Ryden L, on behalf of the ATLAS Study Group: **Low doses versus high doses of the angiotensin converting enzyme inhibitor lisinopril in chronic heart failure: a cost-effectiveness analysis based on the Assessment of Treatment with Lisinopril And Survival (ATLAS) study.** *European Journal of Heart Failure* 2000, **2**:447-454
23. Löthgren M, Zethraeus N: **Definition, interpretation and calculation of cost-effectiveness acceptability curves.** *Health Economics* 2000, **9**:623-630

Pre-publication history
The pre-publication history for this paper can be accessed here:

http://www.biomedcentral.com/1471-2288/2/5/prepub