

# Probabilistic Sensitivity Analysis in Health Economics

Gianluca Baio<sup>1</sup>      A. Philip Dawid<sup>2</sup>

<sup>1</sup>Department of Statistical Science, University College London (UK)

<sup>2</sup>Centre for Mathematical Sciences, University of Cambridge (UK)

## Abstract

Health economic evaluations have recently built upon more advanced statistical decision-theoretic foundations and nowadays it is officially required that uncertainty about both parameters and observable variables be taken into account thoroughly, increasingly often by means of Bayesian methods. Among these, Probabilistic Sensitivity Analysis (PSA) has assumed a predominant role and Cost Effectiveness Acceptability Curves (CEACs) are established as the most important tool. The objective of this paper is to review the problem of health economic assessment from the standpoint of Bayesian statistical decision theory with particular attention to the philosophy underlying the procedures for sensitivity analysis. We advocate here the use of an integrated vision that is based on the value of information analysis, a procedure that is well grounded in the theory of decision under uncertainty, and criticise the indiscriminate use of other approaches to sensitivity analysis.

**Key words:** Bayesian decision theory, Health economic evaluation, Sensitivity analysis.

Research Report No. 292, Department of Statistical Science, University College London

Date: March 2008

## 1 Introduction

Health economic evaluations have recently built upon more advanced statistical decision-theoretic foundations and nowadays it is officially required that uncertainty about both parameters and observable variables be taken into account thoroughly, increasingly often by means of Bayesian methods (Briggs 2001, O’Hagan and Stevens 2001, O’Hagan et al. 2001, Parmigiani 2002*b*, Spiegelhalter and Best 2003, Spiegelhalter et al. 2004). Consequently, sensitivity analysis has become a major part of any new drug approval or reimbursement files and is explicitly required by organisations such as NICE in the UK (Claxton et al. 2005).

This issue has been addressed under two somehow different but complementary perspectives in the health economics literature. On the one hand, much of the work in the late 1990s and in the early years of the new millennium has been devoted to justify sensitivity analysis on the grounds that the so-called “second order uncertainty” (*i.e.* that deriving from the decision maker’s imperfect knowledge of the population parameters) played a major role in the final economic assessment. This was in line with the development of Bayesian methods for cost effectiveness analysis, and produced many specific indicators widely used to report the results of sensitivity analysis.

At the same time, a second point of view was developed more in line with the theory of statistical decision making, which included formally not only an analysis of the cost effectiveness properties of a given intervention, but also a scheme of prioritising future research. In this framework, the uncertainty about the components of the economic model has to be taken into account with reference to the possibility of reviewing the current decision as to which intervention should be selected, in light of new evidence becoming available.

The objective of this paper is to review the problem of health economic assessment from the standpoint of Bayesian statistical decision theory, with specific attention to the philosophy underlying the procedures for sensitivity analysis. In particular, in line with a few recent contributions to the literature, we advocate here the use of an integrated vision that is based on the value of information analysis, a procedure that is well grounded in the theory of decision under uncertainty, and criticise the indiscriminate use of other approaches to sensitivity analysis.

## 2 Framework

We consider a collection of *experimental units*, labelled by  $i$ . We can apply an intervention  $t$  ( $\in \{0, 1\}$ , say) to any unit, and observe a (possibly multivariate) response,  $Y_i$ . The unit  $i$  might itself be a population, and the treatment  $t$  some population-level policy intervention. Sometimes, though not always, the relevant population-level response  $Y_i$  would be an average, or other summary, of individual-level responses within that population.

We wish to decide which treatment to apply to a new unit  $i'$ , judged as

similar, or *exchangeable* (de Finetti 1974), with all the others receiving the same treatment. In other words, the joint subjective distribution of all the  $Y_i$ 's can be regarded as produced as follows. First let a random pair of distributions  $\mathbf{\Pi} = (\Pi^0, \Pi^1)$  have some (subjective) random joint law  $\mathcal{L}$ . Conditionally on the values  $\boldsymbol{\pi} = (\pi^0, \pi^1)$  of  $\mathbf{\Pi}$ , generate each  $Y_i$  independently, with distribution  $\pi^t$  if unit  $i$  received treatment  $t$ .

For the sake of simplicity, we would typically regard the components of the pair  $\mathbf{\Pi}$  as taking values in some parametric family of distribution-pairs of the form  $\mathcal{P} = \{p(\cdot | \theta^t)\}$ , indexed by a vector of variation-independent parameters  $\boldsymbol{\theta} = (\theta^0, \theta^1)$ , where the (subjective) law  $\mathcal{L}$  of  $\mathbf{\Pi}$  can be represented by a joint prior density  $\lambda(\boldsymbol{\theta})$  for the parameter-pair  $\boldsymbol{\Theta} = (\Theta^0, \Theta^1)$  — a similar general framework has also been described by O'Hagan and Stevens (2001).

Note that we can not identify the distribution  $\lambda$  from data. But given extensive empirical evidence, we can discover the realised value  $\boldsymbol{\theta}$  of  $\boldsymbol{\Theta}$ , and thus the realised pair  $\boldsymbol{\pi} = (p(\cdot | \theta^0), p(\cdot | \theta^1))$ . This consideration is instrumental to the procedures of sensitivity analysis to parameters uncertainty — cfr. § 4.

## 2.1 Decision making in health economics

The problem of allocating the best treatment to the new unit  $i'$  is a typical example of decision making under uncertainty, generated by the imperfect knowledge of the random quantities  $(Y, \boldsymbol{\Theta})$ , as well formalised by Parmigiani (2002b) in the medical setting (for a more general discussion, see Raiffa 1968).

We take the standpoint of a governing body who is responsible for guidance on the implementation of alternative interventions for specific public health matters. Typically, a standard programme will be available and a new one is suggested to replace it, perhaps partially or only on specific sub-populations of individuals.

For each possible intervention  $t$ , patient-level data, for instance produced by a set of randomised trials or observational studies, will generally be available in the form  $\mathcal{D}^t = \{y_i : i = 1, \dots, n_t\}$ . We shall generally refer to the whole set of background information as  $\mathcal{D} = \bigcup_t \mathcal{D}^t$ , and for the sake of simplicity consider only the situation where  $t = 0, 1$ .

The overall value of applying treatment  $t$  and obtaining response  $y$  is given by a utility  $u(y, t)$ . A common form of utility is the *net benefit* (Stinnett and Mullahy 1998)

$$u(y, t) = ke - c. \tag{1}$$

Here  $e$  (typically a function of  $y$  only) denotes the realised benefit, suitably measured for instance in terms of *Quality Adjusted Life Years* (QALYs), while  $c$  (which might depend on  $y$  as well as  $t$ ) is the realised cost;  $k$  is a “willingness to pay” parameter used to put cost and benefits on the same scale. The main advantage of the net benefit over other forms of utility function is its linearity in  $(e, c)$ , which allows a simple interpretation and easier calculations.

Notice though that its extensive use in the health economics literature is dictated purely by this pragmatism. In principle, other choices might

be more indicated: for example, the net benefit presupposes that the decision maker is *risk neutral*, which is not always the case in health policy problems (Koerkamp et al. 2007). Therefore, other forms of utility could be used instead, in order to explicitly include a measure of risk aversion.

According to the precepts of Bayesian decision theory and on the basis of the current data  $\mathcal{D}$ , the quality of the future decision  $t$  is the expected utility

$$\begin{aligned}\mathcal{U}^t &:= \mathbb{E}[u(Y, t) \mid \mathcal{D}] \\ &= \int d\theta^t \lambda^t(\theta^t \mid \mathcal{D}) \int dy p(y \mid \theta^t) u(y, t)\end{aligned}\quad (2)$$

where  $\lambda^t(\cdot)$  is the marginal prior (*i.e.* conditional on the background information) density of the random quantity  $\Theta^t$ . The expected utility is obtained averaging over the uncertainty in both *population* (“objective”) and *parameters* (“subjective”) domains.

We would choose the intervention  $t$  to maximise  $\mathcal{U}^t$ : equivalently, we choose  $t = 1$  if (and, henceforth ignoring ties, only if)  $\text{EIB} > 0$ , where

$$\text{EIB} := \mathcal{U}^1 - \mathcal{U}^0$$

is the *expected incremental benefit* (of treatment 1 over treatment 0). Notice that using the net benefit as the utility function, the EIB is more precisely a function of  $k$  (here we consider this dependence only implicitly, though).

The overall utility is obtained as:

$$\begin{aligned}\mathcal{U}^* &:= \max_t \mathcal{U}^t \\ &= \max\{\text{EIB}, 0\} + \mathcal{U}^0.\end{aligned}$$

## 2.2 Cost effectiveness analysis

The objective of health economic evaluations is to compare the population expected performance of the interventions along the two dimensions of interests, *i.e.* the cost and the benefits. In particular, we can define the *increment in mean effectiveness*

$$\Delta_e := \mathbb{E}[e \mid \theta^1] - \mathbb{E}[e \mid \theta^0]$$

and the *increment in mean cost*

$$\Delta_c := \mathbb{E}[c \mid \theta^1] - \mathbb{E}[c \mid \theta^0].$$

When the net benefit is used as the utility function, then the cost effectiveness analysis is focused on the estimation of

$$\text{EIB} = \mathbb{E}[k\Delta_e - \Delta_c] = k\mathbb{E}[\Delta_e] - \mathbb{E}[\Delta_c]$$

where the expectations are now over the subjective distribution of  $\theta$ .

Apparently, any uncertainty over the values of  $\theta$ , and hence of  $(\Delta_e, \Delta_c)$  does not influence the decision process. But is this appropriate?

### 3 Sensitivity Analysis

The definition of the expected utility shows explicitly how, under the Bayesian approach, both individual variations and uncertainty in the value of the parameters are averaged out. From the decision-theoretic point of view, the identification of the overall utility is all that is needed to reach the best decision given the current state of knowledge available to the decision maker. This point has been argued by Claxton (1999b) in the context of health economics.

However, implementing an intervention is typically associated with some risks such as the irreversibility of investments; for this reason, it has been advocated in the literature that health economic evaluations should be subject to some form of Sensitivity Analysis (SA), in order to quantify and qualify the uncertainty underlying the decision process.

Formally, SA is defined in the risk assessment literature as the study of “how uncertainty in some model output can be apportioned, qualitatively or quantitatively, to different sources of uncertainty in the model input” (Saltelli et al. 2004). Parmigiani (2002b) recognises three different forms of SA.

*Marginalisation* is implicit in Bayesian decision-theoretic procedures, such as the one described in § 2.1; the relevant input can be represented by the value of the parameters of the model,  $\Theta$ , whereas the output is the future health-economic outcomes on some reference unit,  $Y_{i'}$ . The uncertainty in all the random quantities is accounted for by the computation of the expected utilities, that are used to determine the optimal decision, but is not analysed separately.

The second form of SA is *Scenario Analysis* (sometimes referred to as *Deterministic Sensitivity Analysis*, DSA). In this case, the experimenter selects a list of interesting values for (some of) the parameters of the model and evaluates the expected outcomes under all these different scenarios. This procedure is easy to implement when the number of parameters involved is relatively small. However, it fails to consider the possible correlation or the underlying uncertainty about the parameters of interest, only focusing on a set of arbitrarily chosen values, regardless on the likelihood of each of them occurring in reality.

These limitations can be overcome by *Probabilistic Sensitivity Analysis* (PSA), a procedure in which all input parameters are considered as random quantities and therefore are associated with a probability distribution that describes the state of science (*i.e.* the background knowledge of the decision maker). This method is clearly in line with the Bayesian analysis, but instead of marginalising out uncertainty in the parameters, PSA is based on a simulation approach (Doubilet et al. 1985). For each iteration  $s = 1, \dots, S$ , a value  $\hat{\theta}_{(s)}$  is simulated from the distribution  $\lambda(\theta | \mathcal{D})$ . The decision analysis is then conducted using that specific value *as if* this were the realised one. The resulting variability in the expected utilities and the influence of each component of  $\theta$  can then be suitably summarised. Much of the recent theoretical work has been devoted specifically to this issue (Briggs et al. 2002, Parmigiani 2002a, O’Hagan et al. 2004, Claxton et al. 2005, Griffin et al. 2006, O’Hagan et al. 2006).

Another reason why PSA plays a major role in health economic evaluation is, as suggested among others by Briggs et al. (2006), that public health or

programme guidance are usually developed as a two-stage, sequential problem. If gathering additional data to supplement the background information  $\mathcal{D}$  is not an option, the decision maker must choose whether to keep the standard programme  $t = 0$ , or to switch to the new one on the basis of some suitable cost-effectiveness measure of utility (*e.g.* the net benefit). According to (2), this evaluation will be based on the predictive distribution of some future outcomes  $y$ , conditionally on  $\mathcal{D}$ .

Conversely, if deferring the final decision in order to gather more data is an available option, then the standard intervention  $t = 0$  will be typically maintained while additional evidence  $\mathcal{E}$  is collected. Once this is available, the analysis can be updated and the utility for each possible intervention will be based on a posterior density of the parameters of interest  $\lambda^t(\theta^t | \mathcal{D}, \mathcal{E})$ , which will induce a predictive distribution for some other future outcomes  $z$  (generally of the same nature as  $y$ ).

Clearly, the option of postponing the decision on cost effectiveness is associated with extra sampling costs. Therefore, it is crucial to assess the impact of this uncertainty on the final decision. A graphical representation of this sequential process is depicted in Figure 1, using a decision tree. In the next two sections we discuss the implementation of PSA according to this framework.

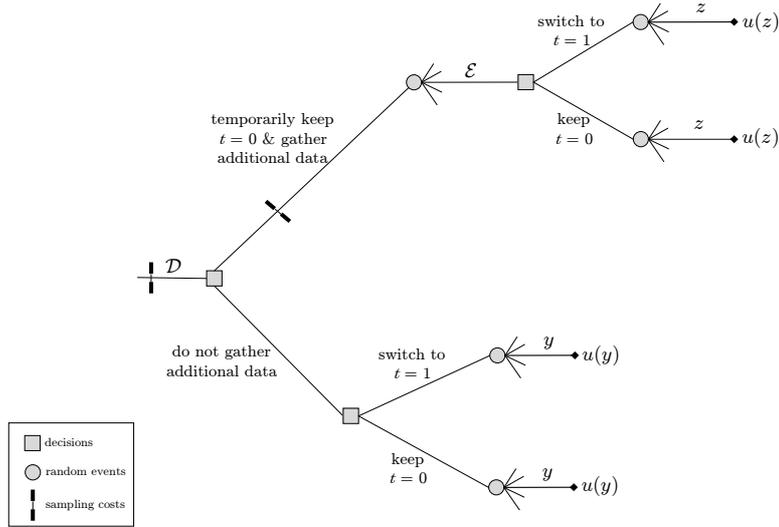


Figure 1: A decision tree representation of the sequential problem of health economic evaluations

## 4 PSA of parameters uncertainty

Let us consider the situation depicted in Figure 1 and assume that, after observing the evidence  $\mathcal{E}$ , the posterior distribution of the parameters is close to

a one-point distribution at the true value: in this case, we shall have effectively learned the realised value  $\theta$  of  $\Theta$ , *i.e.* our uncertainty on the knowledge domain will be resolved.

Now, adopting the intervention  $t$  our expected utility will be  $U(\theta^t)$ , where

$$\begin{aligned} U(\theta) &:= \mathbb{E}[u(Z, t) \mid \mathcal{D}, \mathcal{E}, \theta] \\ &= \int dz p(z \mid \theta) u(z, t). \end{aligned} \quad (3)$$

Consequently, we would choose treatment  $t = 1$  if  $U(\theta^1) > U(\theta^0)$ , or equivalently if  $\text{IB}(\theta) > 0$ , where

$$\text{IB}(\theta) := U(\theta^1) - U(\theta^0) \quad (4)$$

is the *incremental benefit* under parameter-pair  $\theta$ . When the net benefit is used as the utility function,  $\text{IB}(\theta) = k\Delta_e - \Delta_c$ .

The overall “known-distribution” utility is then

$$\begin{aligned} U^*(\theta) &:= \max_t U(\theta^t) \\ &= \max\{\text{IB}(\theta), 0\} + U(\theta^0). \end{aligned}$$

We note also that

$$\mathcal{U}^t = \int d\theta \lambda^t(\theta^t \mid \mathcal{E}) U(\theta^t) \quad \text{and} \quad \text{EIB} = \int d\theta \lambda(\theta \mid \mathcal{E}) \text{IB}(\theta).$$

However, the expectation of the “known distribution” utility calculated with respect to the posterior distribution of the parameters

$$\mathcal{V}^* := \int d\theta \lambda(\theta \mid \mathcal{E}) U^*(\theta)$$

is *not* equivalent to the overall utility  $\mathcal{U}^*$ .

Obviously, in general we shall not be able to learn the value of the parameter  $\Theta$  with certainty and the observation of additional evidence  $\mathcal{E}$  will only be able to reduce our uncertainty on the parameters. Consequently, some indicator is required to adequately describe the difference between the *ideal* and the *actual* decision process.

#### 4.1 PSA in practice: an example

Suppose we set up a simple fictional model of the following form. We assume that the (suitably measured) health economic response  $y = (e, c \mid \theta^t, \mathbf{S}^t) \sim \mathcal{N}(\theta^t, \mathbf{S}^t)$ , where  $\theta^t = (\theta_e^t, \theta_c^t)$  and  $\mathbf{S}^t$  is the population covariance matrix. This is probably the most basic model assuming dependence in  $(e, c)$  — see for instance van Hout et al. (1994). In fact, real-world health economic models are much more complex, but the results that we show throughout the paper using this simple structure are replicable with different distributional assumptions.

Based on some relevant prior knowledge, we can define a subjective prior distribution  $\lambda(\boldsymbol{\theta} \mid \mathcal{D})$  on the parameters, say for the sake of simplicity  $(\boldsymbol{\theta}^0, \boldsymbol{\theta}^1 \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ . Here,  $\boldsymbol{\mu} = (\boldsymbol{\mu}^0, \boldsymbol{\mu}^1)$ , with  $\boldsymbol{\mu}^t = (\mu_e^t, \mu_c^t)$ . Also, for the moment we assume no correlation structure, so that  $\boldsymbol{\Sigma}$  is a diagonal matrix. We suppose further that the values of the hyper-parameters are known, and such that the induced distribution for  $(\Delta_e, \Delta_c)$  has the following properties:  $E(\Delta_c) = \text{sd}(\Delta_c) = \text{£}11,000$ ;  $E(\Delta_e) = \text{sd}(\Delta_e) = 0.5$  QALYs. Obviously, the absence of correlation for the components of  $\boldsymbol{\theta}$  implies that  $\text{Corr}(\Delta_c, \Delta_e) = 0$ . To make this constraints hold, one suitable choice is to set  $\boldsymbol{\mu}^0 = (0.5; 11,000)$ ,  $\boldsymbol{\mu}^1 = (1; 22,000)$  and  $\boldsymbol{\Sigma} = \text{diag}(0.15; 60 \times 10^5; 0.1; 61 \times 10^5)$ .

| Iter                                | Simulated parameters    |                         |                         |                         | Expected utility*                            |                            | Incremental benefit IB( $\boldsymbol{\theta}$ ) |
|-------------------------------------|-------------------------|-------------------------|-------------------------|-------------------------|--|----------------------------|---|
|                                     | $\hat{\theta}_{e(s)}^1$ | $\hat{\theta}_{c(s)}^1$ | $\hat{\theta}_{e(s)}^0$ | $\hat{\theta}_{c(s)}^0$ | $U(\boldsymbol{\theta}^1)$                   | $U(\boldsymbol{\theta}^0)$ |   |
| 1                                   | 0.945                   | 23,100                  | -0.075                  | -5,687                  | 538  | <i>3,797</i>               | -3,259  |
| 2                                   | 0.941                   | 18,720                  | 0.468                   | 10,200                  | <i>4,824</i>                                 | 1,498                      | 3,326   |
| 3                                   | 0.435                   | 14,170                  | -0.537                  | 20,460                  | <i>-3,271</i>                                | -33,900                    | 30,629  |
| 4                                   | 1.484                   | 20,470                  | 0.511                   | 2,717                   | <i>16,640</i>                                | 10,060                     | 6,580   |
| 5                                   | 0.595                   | 17,900                  | 0.911                   | 6,753                   | -3,008                                       | <i>16,030</i>              | -19,038   |
| 6                                   | 0.894                   | 29,400                  | 0.202                   | 19,620                  | <i>-7,036</i>                                | -14,550                    | 7,514   |
| ...                                 |                         |                         | ...                     |                         |  |                            | ...   |
| 1000                                | 1.075                   | 28,770                  | 0.672                   | 24,270                  | <i>-1,903</i>                                | -7,473                     | 5,570   |
| <i>Average over all simulations</i> |                         |                         |                         |                         | <i><math>\mathcal{U}^1 = 3,181.75</math></i> | $\mathcal{U}^0 = 1,398.93$ | EIB = 1,782.82                                  |

\*Obtained for a fixed  $k = \text{£}25,000$ .

Table 1: PSA using simulations. For each iteration and for the overall average, the highest expected utility (*i.e.* the optimal intervention) is typeset in italics

Table 1 shows the PSA derived by the results of 1,000 simulations from this model, using a fixed threshold of willingness to pay,  $k = \text{£}25,000$ . According to the procedure described earlier, we draw a value for each component of  $\boldsymbol{\theta}$  from their joint distribution and we act as if these values were the realised ones, that is as if we had observed further evidence  $\mathcal{E}$  that has resolved our uncertainty on the parameters. At each step, we conduct the decision analysis conditionally on these simulated values to compute the expected utilities (3) and the incremental benefit (4). Averaging over these distributions produces the overall expected values  $\mathcal{U}^t$  and the expected net benefit, EIB. Since its value is positive, the optimal decision is then  $t = 1$ . Figure 2(a) shows the EIB calculated for a range of values for the parameter  $k$ . In addition, using the simulations we can produce a synthesis of the uncertainty underlying this optimal decision.

## 4.2 Cost effectiveness acceptability curves

A very popular indicator commonly used in health economic evaluations to summarise the results of PSA is the *cost effectiveness acceptability curve* (CEAC, van Hout et al. 1994), defined as

$$\text{CEAC} = \Pr(\text{IB}(\boldsymbol{\Theta}) > 0).$$

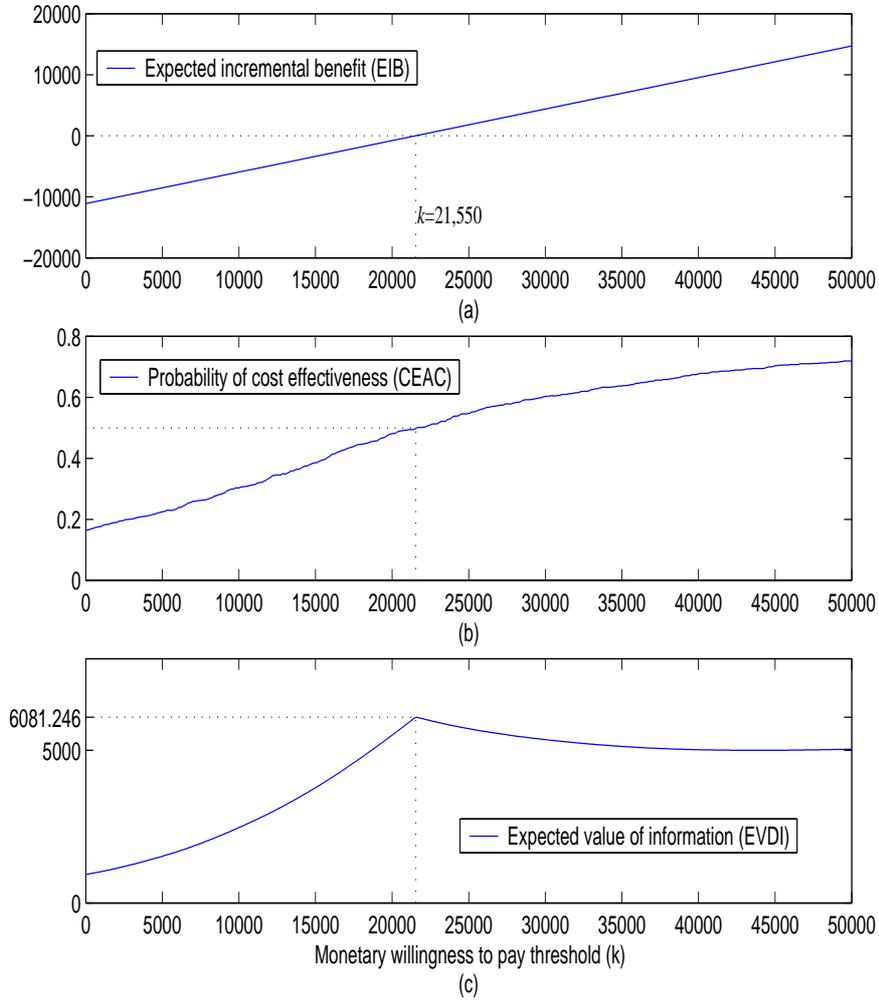


Figure 2: Decision analysis for the simulated data of § 4.1

If the net benefit is used as the utility function, this can be re-expressed as  $CEAC = \Pr(k\Delta_e - \Delta_c > 0)$ , which obviously depends on the willingness to pay parameter  $k$ .

When  $EIB > 0$ , *i.e.* our optimal decision is treatment 1, this is the probability that learning the value of  $\Theta$  (resolving the uncertainty on the parameters) would not change that decision. The rationale of PSA by means of the CEAC is therefore to compare the ideal decision process (by means of the IB) to the

actual one (through the EIB).

By their very nature, CEACs provide a simple synthesis of the uncertainty about the cost effectiveness of a given intervention (Fenwick et al. 2001) and have been widely used in the health economics literature (Briggs 2000, O’Hagan et al. 2000, O’Brien and Briggs 2002, Parmigiani 2002*b*, Spiegelhalter and Best 2003). The main advantage of CEACs is that they simply allow to summarise the probability of cost effectiveness upon varying the willingness to pay parameter, effectively performing a DSA on  $k$ . This circumstance proved to be particularly useful in presenting the results of economic analysis, as decision makers are often not ready to commit to a single value of  $k$  prior to the analysis being performed.

Figure 2(b) shows the CEAC for the fictional example presented above, upon varying the value of the parameter  $k$  in the range £[0; 50, 000]. As is easy to see, for low values of  $k$  in agreement with the region in which  $EIB < 0$ , the probability of cost effectiveness is quite low, indicating higher uncertainty in the decision. In fact, for  $k = £25, 000$  the probability of cost effectiveness is just 0.547. This figure can be deduced from Table 1 as the proportion of simulations for which  $IB(\theta) > 0$ .

Despite their wide use, some critical limitations have been pointed out, the main one being that CEACs do not contain a decision rule. For instance, Felli and Hazen (1999) noticed that they can only address the problem of *how likely* it is that resolving parameters uncertainty will change the optimal decision. Nevertheless, no explicit reference is made to the *possible change in the payoffs*. More recently, Koerkamp et al. (2007) suggested that very different distributions for the IB can produce the same value of the CEAC, which makes it difficult to interpret and might lead to incorrect conclusions for policy makers.

Consequently, by means of CEACs only a partial evaluation of the overall decision process is provided. For this reason, if sensitivity analysis is performed bearing in mind the comprehensive scenario depicted in Figure 1 (*i.e.* with the possibility of reviewing the decision), then the use of CEACs is clearly not ideal.

### 4.3 The value of distributional information

A purely decision-theoretic approach to PSA, overtaking the shortcomings of CEACs, is based on the *value of information* analysis (Howard 1966), an increasingly popular method in health economic evaluations (Felli and Hazen 1998, Felli and Hazen 1999, Claxton 1999*a*, Claxton et al. 2001, Ades et al. 2004, Brennan and Kharroubi 2005, Briggs et al. 2006, Fenwick et al. 2006). In this approach, rather than comparing the IB to the EIB, we seek to evaluate the differences between  $\mathcal{V}^*$ , the expectation of the overall “known distribution” utility (that is obtained when the parameters are assumed known), and  $\mathcal{U}^*$ , the overall utility calculated averaging out uncertainty in both parameters and observable variables.

The value of obtaining the *distributional* information  $\theta$  is

$$VDI(\theta) := U^*(\theta) - \mathcal{U}^*,$$

which represents for each value of  $\theta$  the *opportunity loss* derived by choosing the alternative associated with the highest expected utility  $\mathcal{U}^*$ , instead of the one associated with the highest known distribution utility  $U^*(\theta)$ .

The initial value of the option to learn the parameters before having to decide — the (*prior*) *expected value of distributional information* — is

$$\begin{aligned} \text{EVDI} &:= \int d\theta \lambda(\theta | \mathcal{E}) \text{VDI}(\theta) \\ &= \int d\theta \lambda(\theta | \mathcal{E}) \max\{\text{IB}(\theta), 0\} - \max\{\text{EIB}, 0\}. \end{aligned} \quad (5)$$

This is necessarily non-negative and it places an upper limit to the amount that we would be willing to pay (in utiles) to obtain any information, perfect or imperfect, about  $\Theta$ .

By construction, the EVDI measures the *weighted average opportunity loss* induced by the decision that we make based on the EIB, the weight being the probability of incurring in that loss. Therefore, this measure gives us an appropriately integrated indication of: *a) how much* we are likely to lose if we take the “wrong” decision, and *b) how likely* it is that we take it, as is easily appreciated re-expressing (5) as

$$\begin{aligned} \text{EVDI} &= \text{E}[\max\{\text{IB}(\theta), 0\}] - \max\{\text{EIB}, 0\} \\ &= \text{E}[\text{IB}(\theta) | \text{IB}(\theta) > 0] \times \text{Pr}(\text{IB}(\theta) > 0) - \max\{\text{EIB}, 0\} \\ &= \text{E}[\text{IB}(\theta) | \text{IB}(\theta) > 0] \times \text{CEAC} - \max\{\text{EIB}, 0\} \end{aligned}$$

(the expectations are all taken with respect to the joint distribution of  $\theta$ ).

Figure 2(c) shows the analysis of EVDI as a function of the willingness to pay parameter  $k$  for the fictional example of § 4.1. As one can see, the value of reducing the uncertainty on the parameters increases at a high rate for lower values of the threshold  $k$  (where the probability of changing the decision once the uncertainty on the parameters is resolved is higher — cfr. Figure 2(b)). As appears obvious, when there is lower uncertainty on the actual cost effectiveness of the intervention, gathering new data becomes not as valuable and the EVDI slowly decreases.

Claxton (1999b) argues that if  $\text{EIB} > 0$ , then selecting the treatment  $t = 0$  just because there is a large variability in IB (*i.e.* from the analysis of CEAC) results in imposing unnecessary loss to society as patients can not benefit from a potentially cost effective treatment. On the contrary, the analysis of EVDI provides the decision maker with a rational procedure that allows them to overcome this problem. If the large variability in IB is associated with low cost for additional research, then the decision maker can rightly temporise, or perhaps select the treatment  $t = 1$  only for a subset of the population.

As an example, in the analysis depicted in Figure 2 for  $k \geq \pounds 21,550$  the EIB is positive suggesting that treatment  $t = 1$  should be preferred, even if the CEAC implies a significant degree of riskiness associated with this decision, as the probability of cost effectiveness is only just above 0.5 for  $k = \pounds 21,550$  and

does not exceed 0.7 for the highest willingness to pay threshold considered. The analysis of EVDI allows the decision maker to quantify the resources needed to reduce this riskiness by means of further research, the cost of which can be balanced with the expected returns in terms of increase in the precision of the decision process.

## 5 PSA of parameters uncertainty for instant decisions

In line with the recent developments of the literature, we argued in the previous section that since decision making is concerned with the integration of value and uncertainty, CEACs are inadequate to assess the sensitivity to imperfect knowledge of the parameters, being able to address the latter aspect only.

We now concentrate our attention to situations where the decision maker has to take an instant decision, *i.e.* the option of gathering additional evidence is not formally available. This might be relevant for example in the situation of completely innovative interventions, when it might be unlikely that a new evidence base is made available in the short term. In addition, it might be the case of some reimbursement agencies that are focused only on adoption problems and do not consider explicitly the issue of research prioritising (Sculpher et al. 2006).

Obviously, in this circumstance, the value of future information is not a helpful way to investigate sensitivity in the decision process. Nevertheless, we argue here that also in this case, *regardless on the form of the utility function*, the analysis of CEACs is misleading as it involves aspects that are not actually relevant in terms of decision making.

Consider again the simple multivariate normal model of § 4.1, but now with a correlation structure for the elements of the parameter vector. That is

$$\Sigma = \left( \begin{array}{cc|cc} \tau_e^0 & \phi^0 \sqrt{\tau_e^0 \tau_c^0} & \rho_e \sqrt{\tau_e^0 \tau_e^1} & \psi \sqrt{\tau_e^0 \tau_c^1} \\ \phi^0 \sqrt{\tau_e^0 \tau_c^0} & \tau_c^0 & \xi \sqrt{\tau_c^0 \tau_e^1} & \rho_c \sqrt{\tau_c^0 \tau_c^1} \\ \hline \rho_e \sqrt{\tau_e^1 \tau_e^0} & \xi \sqrt{\tau_e^1 \tau_c^0} & \tau_e^1 & \phi^1 \sqrt{\tau_e^1 \tau_c^1} \\ \psi \sqrt{\tau_c^1 \tau_e^0} & \rho_c \sqrt{\tau_c^1 \tau_c^0} & \phi^1 \sqrt{\tau_e^1 \tau_c^1} & \tau_c^1 \end{array} \right)$$

is the covariance block matrix describing the joint variability of  $\theta_e$  and  $\theta_c$  *within* and *between* the two alternative treatments. In particular, the diagonal blocks of  $\Sigma$  describe the joint variability of the measure of effectiveness and cost within a single treatment regimen. This would only involve the marginal distribution of each parameter  $\Theta^t$ .

On the contrary, the off-diagonal blocks of  $\Sigma$  include:

- the correlations between the measures of effectiveness across different treatments, represented by  $\rho_e$ ;

- the correlations between the measures of cost across different treatments, represented by  $\rho_c$ ;
- the cross correlations between effectiveness and cost across different treatments, represented by  $\psi$  and  $\xi$ ,

*i.e.* they concern the joint distribution of  $(\Theta^0, \Theta^1)$ .

We might reasonably suppose that  $\psi, \xi = 0$ , but since the two programmes are addressing the same disease or condition, it seems plausible to let the average cost and benefits have a certain degree of correlation between them, captured by  $(\rho_e, \rho_c)$ . In practice, parameters independence is often assumed across the different treatments, that is  $\rho_e, \rho_c = 0$  (see for instance the analysis of Spiegelhalter et al. 2004, page 313), but this more general construction shows that we have to be very careful to distinguish the within- and between-population components of variation, even in the case of the simple bivariate normal model.

Considering the usual utility measure  $u(y, t) = ke - c$ , the analysis of § 2.1 suggests that the decision between the alternative treatments is based on the comparison of the expected utilities, which in this case are easily computed as  $\mathcal{U}^t = k\mu_e^t - \mu_c^t$ .

Similarly, (3) is

$$\begin{aligned} U(\theta^t) &= \int de \int dc (ke - c) p(e, c | \theta^t) \\ &= k\theta_e^t - \theta_c^t \end{aligned}$$

and the incremental benefit is easily calculated as

$$\begin{aligned} \text{IB}(\theta) &= (k\theta_e^1 - \theta_c^1) - (k\theta_e^0 - \theta_c^0) \\ &= k(\theta_e^1 - \theta_e^0) - (\theta_c^1 - \theta_c^0) \\ &= k\delta_e - \delta_c, \end{aligned}$$

where  $\delta_e$  and  $\delta_c$  are respectively the realised values of the random variables  $\Delta_e = (\Theta_e^1 - \Theta_e^0)$  and  $\Delta_c = (\Theta_c^1 - \Theta_c^0)$  — cfr. § 2.2.

In this simple case, the random quantity  $\text{IB}(\Theta)$  is a linear combination of normally distributed components and therefore it has a normal distribution with mean

$$m := \text{E}(k\Delta_e - \Delta_c) = k(\mu_e^1 - \mu_e^0) - (\mu_c^1 - \mu_c^0)$$

and variance

$$v^2 := k^2 \text{Var}(\Delta_e) + \text{Var}(\Delta_c) - 2k \text{Cov}(\Delta_e, \Delta_c).$$

Setting  $\psi, \xi = 0$ , but allowing for dependence across treatments and using the properties of variance and covariance of a linear transformation, we also have

$$\alpha := k^2 \text{Var}(\Delta_e) = k^2(\tau_e^1 + \tau_e^0 - 2\rho_e \sqrt{\tau_e^0 \tau_e^1}),$$

$$\beta := \text{Var}(\Delta_c) = \tau_c^1 + \tau_c^0 - 2\rho_c \sqrt{\tau_c^0 \tau_c^1}$$

and

$$\gamma := 2k \text{Cov}(\Delta_e, \Delta_c) = 2k(\phi^1 \sqrt{\tau_e^1 \tau_c^1} - \phi^0 \sqrt{\tau_e^0 \tau_c^0}),$$

from which we can finally derive

$$\text{CEAC} = \Pr(\text{IB}(\boldsymbol{\theta}) > 0) = \Phi\left(\frac{m}{\sqrt{\alpha + \beta - \gamma}}\right),$$

where  $\Phi$  is the Normal cdf.

As is easy to see,  $\alpha$  and  $\beta$  depend on the correlation between  $\boldsymbol{\Theta}^0$  and  $\boldsymbol{\Theta}^1$ , as measured by the parameters  $(\rho_e, \rho_c)$ . This is true irrespective of the form of the utility function. On the contrary, the decision rule is based on  $m$ , a quantity that, regardless on the utility function, only involves the *marginal* distributions of  $\boldsymbol{\Theta}^0$  and  $\boldsymbol{\Theta}^1$ . Notice that due to the linearity of the net benefit, only the mean values  $(\mu_e^t, \mu_c^t)$  are involved in this case, *i.e.* the marginal distributions do not depend on the correlation between the components of the parameters within a single treatment group (described by  $\phi^t$ ). However, different choices for  $u$  could be such that  $\phi^t$  — but not  $(\rho_e, \rho_c)$ ! — becomes relevant for the determination of the expected utility.

This considerations show that PSA should be focused only on the quantities that matter for the decision process and not consider other (irrelevant) aspects of the distribution of the random quantity  $\text{IB}(\boldsymbol{\Theta})$ . Therefore, also in this case the CEAC is not “fit for purpose”. Obviously, if we assume independence across treatments, this problem is not evident. But this analysis shows that in case of instant decisions, marginalisation should be the only criterion; within the Bayesian approach, this would allow to consider the underlying uncertainty in the random quantities, correctly considering the average values of the utilities.

Figure 3 shows the resulting CEAC from three different scenarios about the value of the correlation parameters in the running example. As one can see, despite the fact that the decision process is unique as the marginal averages are the same in all the scenarios, PSA performed by means of CEAC produces three different curves and should therefore be treated with caution.

### 5.1 PSA of the form of the utility function: including a risk aversion parameter in the net benefit

A more appropriate way of conducting a PSA would be to take into account more general forms for the utility function, for instance to include the possibility that the decision maker is not willing to take risks in deciding which intervention is to be implemented. Consider again the same model for  $(e, c \mid \boldsymbol{\theta}^t, \mathbf{S}^t)$ , with

$$\mathbf{S}^t = \mathbf{S} = \begin{pmatrix} \sigma_e^2 & \xi \sigma_e \sigma_c \\ \xi \sigma_e \sigma_c & \sigma_c^2 \end{pmatrix}$$

that is, for the sake of simplicity, we assume here that the elements of the population covariance matrix are common to the two treatment groups (although

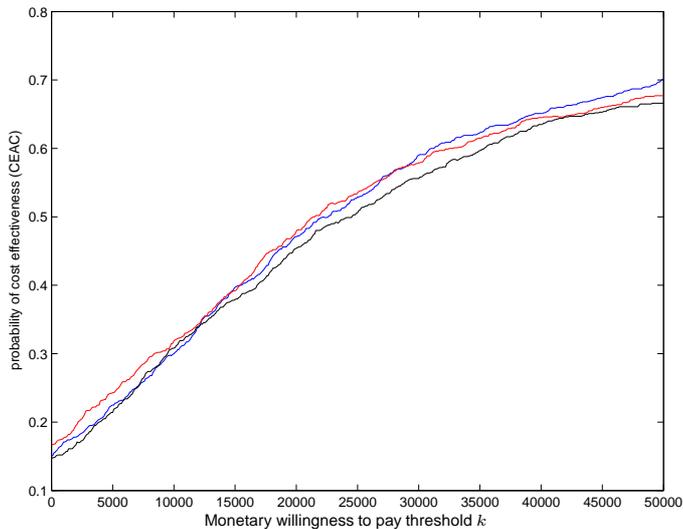


Figure 3: CEAC for the simulated data of § 4.1 upon variations in the correlation parameters. The blue line represents the situation where all the correlation parameters are set to 0; the red line is produced setting  $\phi^0 = 0.4$ ,  $\phi^1 = -0.6$ ,  $\rho_e = -0.8$  and  $\rho_c = 0.2$ ; and finally the black line considers the case where  $\phi^0 = 0.4$ ,  $\phi^1 = -0.6$ ,  $\rho_e = 0.8$  and  $\rho_c = -0.2$

this assumption is not essential). The variable  $y = (ke - c)$  has then a normal distribution  $y \mid \boldsymbol{\theta}^t, \mathbf{S} \sim \mathcal{N}(k\theta_e^t - \theta_c^t, k^2\sigma_e^2 + \sigma_c^2 - 2k\xi\sigma_e\sigma_c)$ .

Instead of the net benefit (1), we now consider a more complex utility function of the form

$$u(y, r, t) = \frac{1}{r} [1 - \exp(-ry)],$$

where  $r > 0$  represents a parameter of *risk aversion*. This represents a generalisation of the standard form of utility including (constant) risk aversion (Raiffa 1968, Jordaan 2005).

Now, in line with the analysis of § 4, the quantity that we should investigate for PSA of uncertainty in the parameters, the known distribution utility (3), is

$$\begin{aligned} U(\boldsymbol{\theta}^t, \mathbf{S}) &= \int dy \frac{1}{r} [1 - \exp(-ry)] p(y \mid \boldsymbol{\theta}^t, \mathbf{S}) \\ &= \frac{1}{r} [1 - M_Y(-r)] \end{aligned}$$

where  $M_Y(-r) := \mathbb{E}[\exp(-rY)]$  is the moment generating function of the random quantity  $Y$  evaluated at the point  $-r$ . Notice that in this case, this expected utility *does* depend on the population variance matrix  $\mathbf{S}$ , as the utility function is no longer linear in  $(e, c)$ . However, thanks to the assumption of

normality, it is easy to calculate

$$U(\boldsymbol{\theta}^t, \mathbf{S}) = \frac{1}{r} \left\{ 1 - \exp \left[ -r (k\theta_e^t - \theta_c^t) + \frac{r^2}{2} (k^2\sigma_e^2 + \sigma_c^2 - 2k\xi\sigma_e\sigma_c) \right] \right\}.$$

Consequently, the incremental benefit (4) is in this case

$$\begin{aligned} \text{IB}(\boldsymbol{\theta}, \mathbf{S}) &= -\frac{1}{r} \exp \left[ \frac{r^2}{2} (k^2\sigma_e^2 + \sigma_c^2 - 2k\xi\sigma_e\sigma_c) \right] \times \\ &\quad \left\{ \exp [-r (k\theta_e^1 - \theta_c^1)] - \exp [-r (k\theta_e^0 - \theta_c^0)] \right\}, \end{aligned}$$

now a function of two deterministic parameters, the willingness to pay  $k$  and the risk aversion  $r$ . Then, it is interesting to summarise the distribution of the IB upon variations in both these quantities.

Figure 4 shows the expected incremental benefit EIB for some different values of  $r$  and as a function of  $k$ . As is possible to see, in this case the EIB is no longer linear; moreover, for low values of  $k$  (*i.e.*  $k < \text{£}22\,000$ ) it is negative regardless on the level of risk aversion. Therefore, in these cases the new treatment would not be enough cost effective for the decision maker.

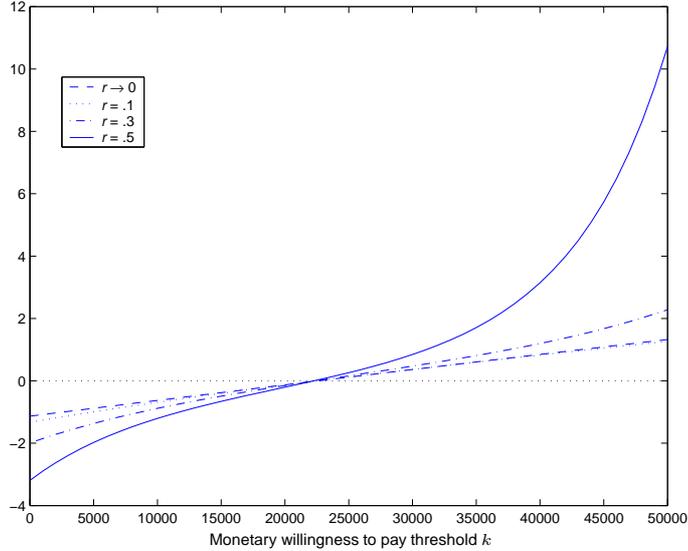


Figure 4: Analysis of the expected incremental benefit including a parameter of risk aversion

Depending on the risk aversion of the decision maker, for values of  $k \in \text{£}[22\,260; 22\,630]$ , the EIB becomes positive, so that the optimal decision is changed and now the new treatment is enough cost effective, increasingly so when the decision maker is more risk averse, as is shown in see Figure 5 (a zoom out of the relevant portion of Figure 4).

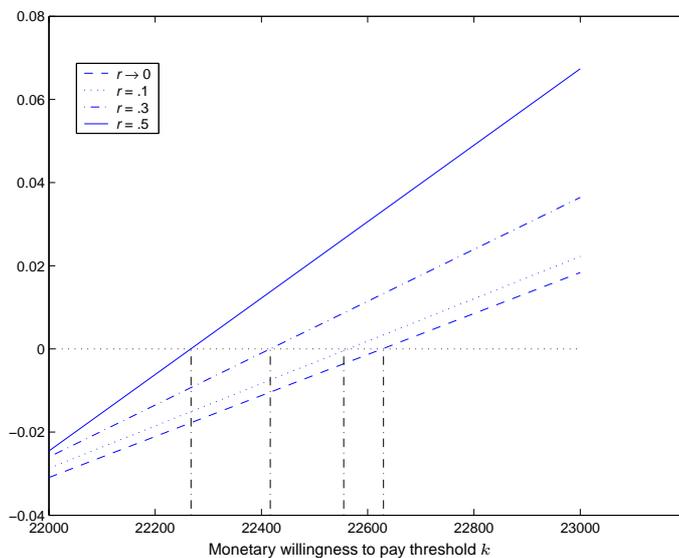


Figure 5: The break even points (*i.e.* the points where the optimal decision is changed from  $t = 0$  to  $t = 1$ ) for different values of  $r$

## 6 Conclusions

In this paper we reviewed the methodology of sensitivity analysis in health economics. The recent past years witnessed the establishment of formal statistical decision-theoretic foundations in this field, along with the increasing awareness of the relevance of monitoring uncertainty in the decision process.

Our standpoint is that PSA is an important component of any health economic analysis; however, we also believe that it should be consistent with the precepts of formal decision-theory, that is it should only concern those aspects that turn out to be crucial in determining the optimal decision. While, on the one hand this concept is not new in the literature, we also argued on the other hand that care should be taken to consider the appropriate context, *e.g.* whether or not further information could be gathered.

In any case, we showed that commonly used methodologies for PSA, such as the analysis of the probability of cost effectiveness summarised by the CEAC, violate the above precepts and therefore we argue that they should be replaced with considerations of expected utility gain, that is the more formal and informative analysis of the value of information, when the option of gathering additional data is available. Moreover, standard methodologies exist that allow the incorporation of risk aversion in the definition of the utility function; these should be exploited to represent more precisely the objective of the public decision makers in problem of identification of an optimal strategy.

## Acknowledgments

The authors are most grateful for the support for this research that was provided by the Leverhulme Foundation and the Economic and Social Research Council to University College London.

## References

- Ades, A., Lu, G. and Claxton, K. (2004), ‘Expected Value of Sample Information Calculations in Medical Decision Modeling’, *Medical Decision Making* **24**, 207–227.
- Brennan, A. and Kharroubi, S. (2005), Efficient Computation of Partial Expected Value of Sample Information Using Bayesian Approximation, Research Report 560/05, Department of Probability and Statistics, University of Sheffield, UK.
- Briggs, A. (2000), ‘Handling Uncertainty in Cost-Effectiveness Models’, *Pharmacoeconomics* **22**, 479–500.
- Briggs, A. (2001), ‘A Bayesian Approach to Stochastic Cost-Effectiveness Analysis’, *International Journal of Technology Assessment in Health Care* **17**, 69–82.
- Briggs, A., Goeree, R., Blackhouse, G. and O’ Brien, B. (2002), ‘Probabilistic Analysis of Cost-Effectiveness Models: Choosing between Treatment Strategies for Gastroesophageal Reflux Disease’, *Medical Decision Making* **4**, 290–308.
- Briggs, A., Schulpher, M. and Claxton, K. (2006), *Decision modelling for health economic evaluation*, Oxford University Press, Oxford, UK.
- Claxton, K. (1999a), ‘Bayesian approaches to the value of information: implications for the regulation of new pharmaceutical’, *Health Economics* **8**, 269–274.
- Claxton, K. (1999b), ‘The irrelevance of inference: a decision-making approach to stochastic evaluation of health care technologies’, *Journal of Health Economics* **18**, 342–364.
- Claxton, K., Neumann, P., Araki, S. and Weinstein, M. (2001), ‘Bayesian Value-Of-Information Analysis’, *International Journal of Technology Assessment in Health Care* **17**, 38–55.
- Claxton, K., Schulpher, M., McCabe, C., Briggs, A., Akehurst, R., Buxton, M., Brazier, J. and O’Hagan, A. (2005), ‘Probabilistic sensitivity analysis for NICE technology assessment: not an optional extra’, *Health Economics* **14**, 339–347.

- de Finetti, B. (1974), *Theory of Probability*, Vol. 1, John Wiley and Sons, New York, NY.
- Doubilet, P., Begg, C., Weinstein, M., Braun, P. and McNeil, B. (1985), ‘Probabilistic sensitivity analysis using Monte Carlo simulation. A practical approach’, *Medical Decision Making* **5**, 157–177.
- Felli, J. and Hazen, G. (1998), ‘Sensitivity analysis and the expected value of perfect information’, *Medical Decision Making* **18**, 95–109.
- Felli, J. and Hazen, G. (1999), ‘A Bayesian approach to sensitivity analysis’, *Health Economics* **8**, 263–268.
- Fenwick, E., Claxton, K. and Schulpher, M. (2001), ‘Representing uncertainty: the role of cost effectiveness acceptability curves’, *Health Economics* **10**, 779–787.
- Fenwick, E., Palmer, S., Claxton, K., Sculpher, M., Abrams, K. and Sutton, A. (2006), ‘An Iterative Bayesian Approach to Health Technology Assessment: Application to a Policy of Preoperative Optimization for Patients Undergoing Major Elective Surgery’, *Medical Decision Making* **26**, 480–496.
- Griffin, S., Claxton, K., Hawkins, N. and Sculpher, M. (2006), ‘Probabilistic Analysis and Computationally Expensive Models: Necessary and Required?’, *Value in Health* **9**, 244–252.
- Howard, R. (1966), Information Value Theory, in ‘IEEE Transactions on System Science and Cybernetics’, SCC-2, (1) 22-26.
- Jordaan, I. (2005), *Decisions under uncertainty*, Cambridge University Press, Cambridge, UK.
- Koerkamp, B., Hunink, M., Stijnen, T., Hammitt, J., Kuntz, K. and Weinstein, M. (2007), ‘Limitations of acceptability curves for presenting uncertainty in cost-effectiveness analyses’, *Medical Decision Making* **27** (2), 101–111.
- O’Brien, B. and Briggs, A. (2002), ‘Analysis of uncertainty in health care cost-effectiveness studies: an introduction to statistical issues and methods’, *Statistical Methods in Medical Research* **11**, 455–468.
- O’Hagan, A., McCabe, C., Akehurst, R., Brennan, A., Briggs, A., Claxton, K., Fenwick, E., Fryback, D., Sculpher, M., Spiegelhalter, D. and Willan, A. (2004), ‘Incorporation of uncertainty in health economic modelling studies’, *Pharmacoeconomics* **23**, 539–536.
- O’Hagan, A. and Stevens, J. (2001), ‘A framework for cost-effectiveness analysis from clinical trial data’, *Health Economics* **10**, 303–315.
- O’Hagan, A., Stevens, J. and Montmartin, J. (2000), ‘Inference for the cost-effectiveness acceptability curve and cost-effectiveness ratio’, *Pharmacoeconomics* **17**, 339–349.

- O'Hagan, A., Stevens, J. and Montmartin, J. (2001), 'Bayesian cost effectiveness analysis from clinical trial data', *Statistics in Medicine* **20**, 733–753.
- O'Hagan, A., Stevenson, M. and Madan, J. (2006), 'Monte Carlo probabilistic sensitivity analysis for patient level simulation models: efficient estimation of mean and variance using ANOVA', *Health Economics* .
- Parmigiani, G. (2002a), 'Measuring uncertainty in complex decision analysis models', *Statistical Methods in Medical Research* **11**, 513–537.
- Parmigiani, G. (2002b), *Modeling in Medical Decision Making*, John Wiley and Sons, New York, NY.
- Raiffa, H. (1968), *Decision Analysis – Introductory Lectures on Choices under Uncertainty*, Addison Wesley, Reading, MA.
- Saltelli, A., Tarantola, S., Campolongo, F. and Ratto, M. (2004), *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models*, John Wiley and Sons, Chichester, UK.
- Sculpher, M., Claxton, K., Drummond, M. and McCabe, C. (2006), 'Whither trial-based economic evaluation for health decision making?', *Health Economics* **15**, 677–687.
- Spiegelhalter, D., Abrams, K. and Myles, J. (2004), *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*, John Wiley and Sons, Chichester, UK.
- Spiegelhalter, D. and Best, N. (2003), 'Bayesian approaches to multiple sources of evidence and uncertainty in complex cost-effectiveness modelling', *Statistics in Medicine* **22**, 3687–3709.
- Stinnett, A. and Mullahy, J. (1998), 'Net health benefits: a new framework for the analysis of uncertainty in cost effectiveness analysis', *Medical Decision Making* **18 (Suppl)**, S68–S80.
- van Hout, B., Gordon, M. and Rutten, F. (1994), 'Costs, effects and C/E ratios alongside a clinical trial', *Health Economics* **3**, 309–319.