

A Fully Bayesian Cost–Effectiveness Analysis using Conditionally Specified Prior Distributions

M. Martel, M.A. Negrín and F.J. Vázquez–Polo

Department of Quantitative Methods.

University of Las Palmas de G.C.

Campus de Tafira.

35017 Las Palmas de Gran Canaria, Spain.

{mmartel or mnegrin or fjvpolo}@dmc.ulpgc.es

Abstract

The Bayesian approach to statistics has been growing rapidly in popularity as an alternative to the classical approach in the economic evaluation of health technologies, due to the significant benefits it affords. One of the most important advantages of Bayesian methods is their incorporation of prior information. Thus, use is made of a greater amount of information, and so stronger results are obtained than with frequentist methods.

In a cost-effectiveness analysis, we relate the costs and effectiveness of the two technologies being compared, the parameters of interest being the mean effectiveness and mean cost of each. The most common prior structure for these two parameters is the bivariate normal structure. Since Stevens and O’Hagan (2002) showed that the elicitation of a prior distribution on the parameters of interest plays a crucial role in a Bayesian cost-effectiveness analysis, relatively few papers have addressed this issue, although Leal et al. (2007) recently presented a computer-based model to elicit uncertainty on parameters.

In this paper we study the use of a more general (and flexible) family of prior distributions for the parameters. In particular, we assume that the conditional densities of the parameters are all normal. This structure allows us to incorporate a large range of prior information. The bivariate normal distribution is included as a particular case of the conditional prior structure.

KEY WORDS: *Bayesian analysis, cost–effectiveness, prior information, conditionally specified distributions.*

1 Introduction

Clinical research is fundamentally a dynamic process in which any study must be considered in the context of continual updating of the state of the art. The Bayesian method is of a dynamic nature in which initial beliefs, determined on the basis of a prior distribution, are modified by new data, using the Bayes theorem.

Spiegelhalter et al. (1994) and Jones (1996) were the first to discuss the Bayesian approach for statistical inference in the comparison of health technologies. Since then, many studies have proposed the Bayesian approach to compare treatment options by means of cost-effectiveness analysis (Brophy and Joseph, 1995; Heitjan, 1997; Al and Van Hout, 2000; Fryback et al., 2001; Vanness and Kim, 2002; Chilcott et al., 2003; Stevens et al., 2002; among others).

To compare the results of two treatments we calculate the most prevalent measures of the cost-effectiveness analysis of a new treatment: the incremental cost-effectiveness ratio (ICER), the incremental net benefit (INB) and the cost-effectiveness acceptability curve (CEAC).

The ICER is defined by:

$$ICER = \frac{\gamma_1 - \gamma_0}{\mu_1 - \mu_0} = \frac{\Delta\gamma}{\Delta\mu}, \quad (1)$$

where γ_j and μ_j are the average costs and efficacies under treatment j (1, new; and 0, for the current or control treatment), respectively.

The INB of treatment 1 versus treatment 0 is defined as

$$INB(R_c) = R_c \cdot (\Delta\mu) - (\Delta\gamma), \quad (2)$$

for each R_c , which is interpreted by O'Hagan and Stevens (2001) as the cost that decision-takers are willing to accept in order to increase the effectiveness of the treatment applied by one unit. Thus, analyzing whether the alternative treatment is more cost effective than the control treatment is equivalent to determining whether $INB(R_c)$ is positive.

In practice, it is not a simple matter for the decision-taker to determine a single R_c , and so a cost-effectiveness acceptability curve (CEAC) is constructed. This curve provides a graphical representation of the probability of the alternative treatment being preferred ($\Pr(INB(R_c) > 0)$) for each value R_c . This interpretation of the CEAC, in terms of probability, is only possible when the Bayesian approach is adopted.

Most published studies on cost-effectiveness analysis assume normality of the cost and effectiveness generation distribution (Willan and O'Brien, 1996; Laska et al., 1997; Stinnett and Mullahy, 1998; Tambour et al., 1998; Heitjan et al., 1999). The normal-case was examined by O'Hagan et al. (2001), who considered the patient level-data $\{x_{ij} : i = 1, 2, \dots, n_j; j = 0, 1\}$ from a clinical trial, where $x_{ij} = (e_{ij}, c_{ij})$. The index j is used to note the treatment and n_j denotes the sample size for each treatment j .

We denote by $f(x_{ij}|\theta_j)$ the parametric distribution generating data x_{ij} from treatment j . The parameters of this function are the mean cost (γ_j), the mean efficacy (μ_j) and the

variance–covariance matrix Σ_j . Then the likelihood is:

$$\ell(\bar{x}|\theta_0, \theta_1) = \prod_{j=0}^1 \prod_{i=1}^{n_j} f(x_{ij}|\theta_j). \quad (3)$$

The most usual model for data \bar{x} is to assume that $f(x_{ij}|\theta_j)$ is a bivariate normal distribution for each treatment j

$$f(x_{ij}|\theta_j) = (2\pi|\Sigma_j|)^{-1/2} \exp\left\{-\frac{1}{2}(x_{ij} - \alpha_j)' \Sigma_j^{-1} (x_{ij} - \alpha_j)\right\}. \quad (4)$$

A Bayesian analysis of model (4) requires the specification of a prior distribution on $(\theta_0, \theta_1) = ((\alpha_0, \Sigma_0), (\alpha_1, \Sigma_1))$, where $\alpha_j = (\mu_j, \gamma_j)$, i.e. the mean efficacy and cost for treatment i , respectively.

Advocates of frequentist statistical methods argue that prior information is intrinsically subjective and therefore has no place in science. However, the incorporation of prior information allows Bayesian methods to access more information and so to produce stronger inferences. The incorporation of prior information can provide more realistic conclusions, particularly where sample sizes are relatively small, as is often the case in cost–effectiveness analysis. Stevens and O’Hagan (2002) discuss the advantages of incorporating prior information in cost–effectiveness analysis of clinical trial data, exploring mechanisms to safeguard scientific rigor in the use of prior information.

A convenient class of prior distributions is a general conditional–conjugate prior. Specifically, O’Hagan et al. (2001) assume a bivariate normal distribution for α_j and an inverse Wishart prior distribution for the variance matrices Σ_j .

In this paper we study the use of a more general family of prior distributions for the parameters α_j . In particular, we assume that the conditional density of μ_j for a given γ_j and the conditional density of γ_j for a given μ_j are both normal. This assumption varies distinctly from one of classical bivariate normality with its familiar elliptical contours.

The paper is organized as follows. Section 2 presents the normal case of the cost–effectiveness analysis with prior distributions based on conditional specification. In Section 3 some examples are given to show that the methodology is readily applicable. We use a practical application with real data from a clinical trial, comparing two alternative treatments for asymptomatic HIV patients. Section 4 presents a discussion of the results obtained and some conclusions are drawn.

2 Bayesian cost–effectiveness analysis with prior distributions based on conditional specification

Our basic prior formulation for model (4) assumes that the joint distribution factorizes as

$$\pi(\theta) = \pi(\alpha_0) \cdot \pi(\alpha_1) \cdot \pi(\Sigma_0) \cdot \pi(\Sigma_1). \quad (5)$$

That is, we assume independence between treatments and between the means (α_j) and the variance matrix (Σ_j). We assume inverse Wishart prior distributions for the variance matrices Σ_0 and Σ_1 . Specifically, we take $\Sigma_j \sim IW(A_j, f_j)$ the prior density of which is

$$\pi(\Sigma_j) \propto |\Sigma_j|^{-(f_j+3)/2} \exp \left\{ \text{tr}(\Sigma_j^{-1} A_j)/2 \right\},$$

over the space of positive–definite 2×2 matrices. Thus f_j is the prior degrees of freedom parameter and the prior expectation of Σ_j is $(f_j - 3)^{-1} A_j$, provided $f_j > 3$.

Likewise it is reasonable to assume a prior normal distribution of μ_j for a given γ_j and a prior normal distribution of γ_j for a given μ_j . The bivariate normal distribution was proposed in the paper of O’Hagan et al. (2001), but that is only a particular case with normal conditionals. In this paper we extend the analysis to consider the class of all joint densities with normal conditionals.

Castillo and Galambos (1989) show the specification of the class of all bivariate densities with normal conditionals. We seek to obtain all joint densities $f_{\mu,\gamma}(\mu, \gamma)$ such that every conditional density of μ given $\gamma = \gamma^0$ is normal with mean $\delta_1(\gamma)$ and variance $\sigma_1^2(\gamma)$ (which may depend on γ) and every conditional density of μ given $\mu = \mu^0$ with mean $\delta_2(\mu)$ and variance $\sigma_2^2(\mu)$ (which may depend on μ).

These authors found that all the bivariate densities with normal conditionals are those of the form

$$f_{\mu,\gamma}(\mu, \gamma) = \exp \left\{ (1, \mu, \mu^2) \begin{pmatrix} m_{00} & m_{01} & m_{02} \\ m_{10} & m_{11} & m_{12} \\ m_{20} & m_{21} & m_{22} \end{pmatrix} \begin{pmatrix} 1 \\ \gamma \\ \gamma^2 \end{pmatrix} \right\}. \quad (6)$$

The conditional expectations and variances are:

$$E[\mu|\gamma] = -\frac{m_{10} + m_{11} \cdot \gamma + m_{12} \cdot \gamma^2}{2(m_{20} + m_{21} \cdot \gamma + m_{22} \cdot \gamma^2)}, \quad (7)$$

$$\text{Var}[\mu|\gamma] = -\frac{1}{2(m_{20} + m_{21} \cdot \gamma + m_{22} \cdot \gamma^2)},$$

$$E[\gamma|\mu] = -\frac{m_{01} + m_{11} \cdot \mu + m_{21} \cdot \mu^2}{2(m_{02} + m_{12} \cdot \mu + m_{22} \cdot \mu^2)},$$

$$\text{Var}[\gamma|\mu] = -\frac{1}{2(m_{02} + m_{12} \cdot \mu + m_{22} \cdot \mu^2)}.$$

The distribution with density of the form (6) is an eight–parameter family of densities. The coefficient m_{00} is a normalizing constant that is determined by the other coefficients m and the requirement that the density integrates to 1.

We encounter a great variety of distributions for different values of the m parameters. Some of these distributions are markedly different from classical bivariate normal densities. We now show the values of the m parameters for some particular cases.

- **Independence:** If we assume prior independence between the mean of the effectiveness (μ) and the mean of the costs (γ) for a given treatment, the conditional distributions do not depend on the other parameter, and the conditional expectations and variances will be of the form:

$$\begin{aligned} E[\mu|\gamma] &= E[\mu] = -\frac{m_{10}}{2 \cdot m_{20}}, \\ \text{Var}[\mu|\gamma] &= \text{Var}[\mu] = -\frac{1}{2 \cdot m_{20}}, \\ E[\gamma|\mu] &= E[\gamma] = -\frac{m_{01}}{2 \cdot m_{02}}, \\ \text{Var}[\gamma|\mu] &= \text{Var}[\gamma] = -\frac{1}{2 \cdot m_{02}}. \end{aligned} \quad (8)$$

Thus, the conditions for independence are that the m 's satisfy the following conditions:

$$m_{11} = m_{12} = m_{21} = m_{22} = 0, \quad m_{20} < 0, \quad m_{02} < 0. \quad (9)$$

- **Bivariate normal distribution**

The terms μ and γ are said to have a bivariate normal distribution, denoted by $(\mu, \gamma) \sim \mathcal{N}_2(\delta_\mu, \delta_\gamma, \sigma_\mu, \sigma_\gamma, \rho)$, if

$$\pi(\mu, \gamma | \delta_\mu, \delta_\gamma, \sigma_\mu, \sigma_\gamma, \rho) = \frac{1}{2\pi\sigma_\mu\sigma_\gamma\sqrt{1-\rho^2}} \exp\left\{\frac{Q}{2(1-\rho^2)}\right\},$$

where Q is the quadratic expression

$$Q = \frac{(\mu - \delta_\mu)^2}{\sigma_\mu^2} - \frac{2\rho(\mu - \delta_\mu)(\gamma - \delta_\gamma)}{\sigma_\mu\sigma_\gamma} + \frac{(\gamma - \delta_\gamma)^2}{\sigma_\gamma^2}.$$

The conditional distributions, too, are normal with mean and variance

$$\begin{aligned} E(\mu|\gamma) &= \delta_\mu + \frac{\rho\sigma_\mu}{\sigma_\gamma}(\gamma - \delta_\gamma), \\ \text{Var}(\mu|\gamma) &= \sigma_\mu^2(1 - \rho^2), \\ E(\gamma|\mu) &= \delta_\gamma + \frac{\rho\sigma_\gamma}{\sigma_\mu}(\mu - \delta_\mu), \\ \text{Var}(\gamma|\mu) &= \sigma_\gamma^2(1 - \rho^2). \end{aligned} \quad (10)$$

From expressions (7) and (10) we can elicit the prior information. Thus, the condition for the bivariate normal distribution is that the m 's satisfy the following conditions (Arnold et al., 2001a,b).

$$m_{12} = m_{21} = m_{22} = 0, \quad m_{20} < 0, \quad m_{02} < 0 \quad \text{and} \quad m_{11}^2 < 4m_{02}m_{20}. \quad (11)$$

- **Other cases:**

The improvement obtained from the use of conditionally specified priors is the wide range of prior information that may thus be elicited. Besides the bivariate normal distribution, there are other combinations of m 's that have non-normal marginal densities. In particular, we can encounter bimodal or even trimodal densities. These distributions must satisfy the sufficient conditions for integrability of (6) (Gelman and Meng, 1991; Arnold et al., 2000; Arnold et al., 2001a).

$$m_{22} < 0, \quad 4m_{22}m_{02} > m_{12}^2, \quad 4m_{22}m_{20} > m_{21}^2. \quad (12)$$

However, we must pay a price for the flexibility of our prior structure, namely that there are eight hyperparameters to assess. We recommend the method for matching conditional moments proposed by Arnold et al. (1998). For a conditionally specified prior such as (6-7), we can try to match conditional moments, whose approximate values will be supplied by the experts. In our analysis, we need at least eight conditional moments to determine all the hyperparameters. However, eight conditional moments might not be enough to determine the prior information. It is preferable for the expert to supply more than eight conditional moments. Moreover, it is likely that such prior values will not be consistent. Therefore, we can elicit the eight hyperparameters by minimizing the deviance between the conditional moments of the form (7) and those provided a priori by the experts.

Let us assume that prior assessed values for the conditional means and variances of the effectiveness and cost are obtained for several different given values of the cost and effectiveness, respectively.

$$E[\mu|\gamma_{p_1}] = e_{p_1} \quad \forall p_1 = 1, 2, \dots, P_1. \quad (13)$$

$$\text{Var}[\mu|\gamma_{p_2}] = \text{var}(e)_{p_2} \quad \forall p_2 = 1, 2, \dots, P_2.$$

$$E[\gamma|\mu_{p_3}] = c_{p_3} \quad \forall p_3 = 1, 2, \dots, P_3.$$

$$\text{Var}[\gamma|\mu_{p_4}] = \text{var}(c)_{p_4} \quad \forall p_4 = 1, 2, \dots, P_4.$$

where $P_1 + P_2 + P_3 + P_4 \geq 8$.

A unique solution for this system of equations is unlikely to be possible for any choice of the eight hyperparameters. One solution is to allow any deviance between the prior conditional moment and the knowledge of the expert. We define as the objective function the sum of the squared deviances. The hyperparameters are obtained by minimizing the objective function subject to constraints (12). The prior distribution obtained must be checked by the experts so as not to obtain local minima in the optimization.

Prior information that cannot be specified using the bivariate normal distribution can be elicited by this method. We show some examples in which the information should be specified on the basis of conditionally specified prior distributions.

Assume we have prior information about the effectiveness of a given treatment from two different studies. Suppose, moreover, that the effectiveness and the cost estimated for each treatment are different, and that both studies are equally credible. The bivariate normal distribution is not appropriate to describe this prior information, but a conditionally specified prior distribution can incorporate the bimodality of this prior information.

Let us now assume there is prior information about the effectiveness and cost of a treatment (mean and variance). Moreover, we know the treatment involves some risk, and that there is a possibility that complications may appear. In this case, the effectiveness is lower and the costs higher. This prior information can be modelled by a bimodal bivariate distribution, using conditionally specified prior distributions.

3 Example

The data used in this section were obtained from a real clinical trial in which a comparison was made between two highly–active antiretroviral treatment protocols applied to asymptomatic HIV patients (Pinto et al., 2000).

We obtained data on the direct costs (of drugs, medical visits and diagnostic tests), and on the effectiveness, based on health–related life–quality variables, using EuroQol–5D. This is an instrument for the self–evaluation of personal health, consisting of five questions that investigate five aspects of health–related life quality, based on a visual analogue scale (VAS) (Brooks, 1996). This scale simulates a thermometer with a minimum of 0 and a maximum of 100. The 0 represents the worst health state imaginable, and the 100, the best. As a measure of effectiveness, we used the variation in the VAS at the end of the study.

In this exercise, we compared two three–way treatment protocols. The first of these (d4T + 3TC + IND) combined the drugs estavudine (d4T), lamivudine (3TC) and indinavir (IND); the second treatment protocol (d4T + ddl + IND) combined estavudine (d4T), didanosine (ddl) and indinavir (IND).

Table 1 summarizes the statistical data. The d4T + ddl + IND treatment was more costly than the d4T + 3TC + IND treatment, by an average of 164.82 euros. When the VAS variation was used as the measure of effectiveness, the d4T + ddl + IND treatment was more effective because, on average, the patients who received this treatment experienced an improvement in their life quality of 4.94 units, while those who were given the d4T + 3TC + IND treatment only experienced a VAS improvement of 4.56 units.

- Independence

The first analysis shown is made under the assumption of independence. For the purpose of this analysis, we take the design of the study to imply prior expectations

Table 1: Statistical summary of costs (in thousands of euros) and effectiveness (change in VAS).

Statistical measure	d4T + 3TC + IND		d4T + ddl + IND	
	Cost	Change in VAS	Cost	Change in VAS
Mean	7.142	4.56	7.307	4.94
s.d.	0.001573	15.17	0.001720	13.98
n	$n_0 = 268$		$n_1 = 93$	

for the parameters of interest. We assume an average of 4.5 units of effectiveness for the control treatment (d4T + 3TC + IND), with a prior variance of 2.25. For this treatment, the design anticipates an average cost of 5 thousand euros, with a variance of 4. The value of the m parameters is elicited directly, in the knowledge of the prior mean and variance of effectiveness and cost. For this prior information, the values are:

$$m_{01} = 1.25, m_{02} = -0.125, m_{10} = 2, m_{11} = 0,$$

$$m_{12} = 0, m_{20} = -0.2222, m_{21} = 0, m_{22} = 0.$$

The elicitation process is very similar for the new treatment (d4T + ddl + IND). In this case, we assume that the treatment is less effective, with an average of 4 units of effectiveness and a prior variance of 2.5. This treatment is also more expensive, with a prior mean cost of 6000 euros, and a variance of 6.25. The values of the m parameters for this treatment are

$$m_{01} = 0.96, m_{02} = -0.08, m_{10} = 1.6, m_{11} = 0,$$

$$m_{12} = 0, m_{20} = -0.2, m_{21} = 0, m_{22} = 0.$$

We also use a diffuse prior distribution for the matrix variances Σ_i . Under the assumption of noninformative priors, we set $A_0 = A_1 = \text{diag}(1, 1)$, $f_0 = f_1 = 2$, where $\text{diag}(a_i)$ is the $n \times n$ diagonal matrix with a_i elements. This assumption is repeated in the following analysis.

Figure 1 shows the joint distribution of the prior information of effectiveness and cost for each treatment, and the joint distribution of the prior incremental effectiveness and cost between treatments.

The posterior distribution has been simulated using WinBUGS (Spiegelhalter et al., 2000). Three parallel chains and a single long chain were used for diagnostic assessment (checked using CODA software). A total of 10000 iterations were carried out

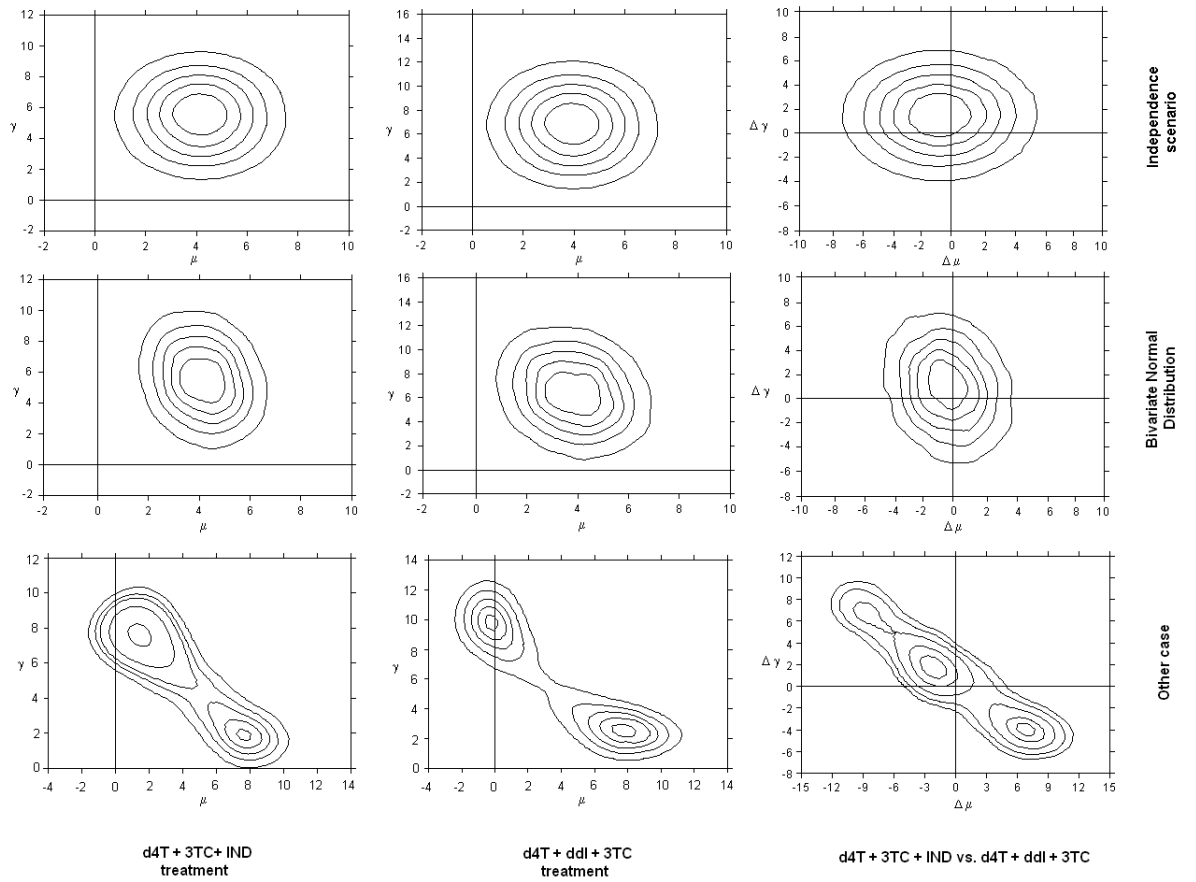


Figure 1: Prior joint distribution of μ and γ .

(after a burn–in period of 40000 simulations). The constant m_{00} is not required to ensure convergence.

Table 2 shows the posterior analysis for the independence case. The posterior incremental effectiveness is estimated as -0.02928 units with a standard deviation of 1.328. The incremental cost is estimated as 0.162 units.

Figure 2 shows the cost–effectiveness plane and the cost–effectiveness acceptability curve; it is apparent that the treatment (d4T + ddl + IND) will never be preferable to the treatment (d4T + 3TC + IND), as the probability of a positive INB is always below 50%.

Table 2: Posterior moments: mean and standard deviation.

	Independence	Bivariate-Normal distribution	Bimodal case
μ_0	4.535 (0.7905)	4.443 (0.786)	3.345 (0.7774)
γ_0	7.138 (0.09529)	7.138 (0.0953)	7.13 (0.09561)
μ_1	4.506 (1.069)	4.42 (1.059)	2.125 (0.8673)
γ_1	7.3 (0.1785)	7.299 (0.1785)	7.292 (0.1829)
$\Delta\mu$	-0.02928 (1.328)	-0.02364 (1.317)	-1.22 (1.165)
$\Delta\gamma$	0.162 (0.2021)	0.1612 (0.2021)	0.162 (0.2064)

- Bivariate normal distribution:

The previous analysis was repeated under the assumption of a correlation between cost and effectiveness for each treatment. In particular, we assumed a prior correlation of $\rho = -0.2$. By incorporating this information to the prior information described in the previous subsection, we elicited the following prior structure:

$$m_{01} = 1.614583333, m_{02} = -0.1302083333, m_{10} = 2.4305555, m_{11} = -0.0694444,$$

$$m_{12} = 0, m_{20} = -0.23148148, m_{21} = 0, m_{22} = 0,$$

for the (d4T + 3TC + IND) treatment, and

$$m_{01} = 1.210818510, m_{02} = -0.08333333, m_{10} = 1.982894432, m_{11} = -0.05270462766,$$

$$m_{12} = 0, m_{20} = -0.2083333333, m_{21} = 0, m_{22} = 0,$$

for the (d4T + ddl + IND) treatment.

Figure 1 shows the joint distribution of the prior information of effectiveness and cost for each treatment, and the joint distribution of the prior incremental effectiveness and cost between treatments.

The results are very similar to those of the previous analysis. Figure 2 shows the cost–effectiveness plane and the cost–effectiveness acceptability curve.

- Other case:

In this example, we show a prior bimodal density for the effectiveness and cost for each treatment.

Suppose that, if there are no complications during the study, the mean effectiveness of the (d4T + 3TC + IND) treatment is close to 8 units, and the mean cost is about

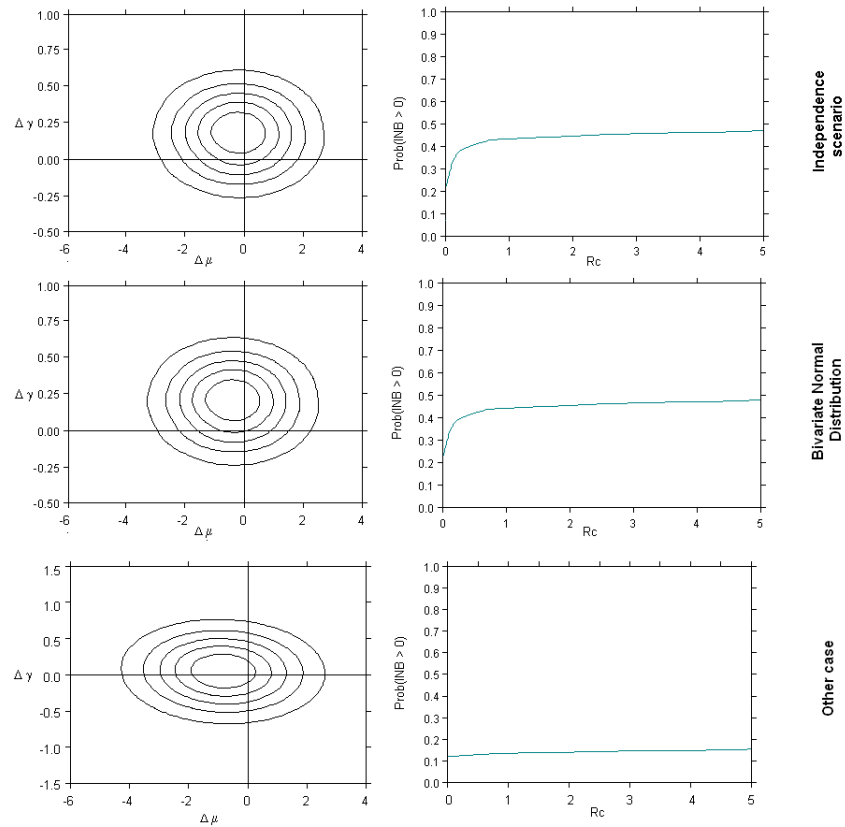


Figure 2: Posterior joint distribution of μ and γ .

2000 euros. However, there is a certain probability that treatment complications will appear. With complications, the mean effectiveness is close to 2 units, and the mean cost is 8000 euros.

To elicit this prior information through a conditionally specified prior distribution, we compile information about conditional moments. Table 3 shows the conditional moments employed in the elicitation process.

With respect to the optimization problem explained in the previous section, we now calculate the values of the hyperparameters.

$$m_{01} = 9.393073, m_{02} = -0.6198139, m_{10} = 8.144171, m_{11} = -1.811442,$$

Table 3: Prior conditional moments.

Moment	Condition			Condition		
		(d4T+3TC+IND)	(d4T+ddl+IND)		(d4T+3TC+IND)	(d4T+ddl+IND)
E($\mu \gamma$)	$\gamma = 1$	9	9	$\gamma = 1.5$	8.2	8.2
	$\gamma = 2$	8	8	$\gamma = 2.5$	8	8
	$\gamma = 3$	7.5	7.5	$\gamma = 3.5$	7.5	7.5
	$\gamma = 4$	7.2	7.2	$\gamma = 5$	3.2	3.2
	$\gamma = 6$	2.2	2.2	$\gamma = 7$	0.7	0.7
	$\gamma = 8$	0.7	0.7	$\gamma = 9$	0	
	$\gamma = 9.5$	0	0	$\gamma = 10$	0.5	-0.5
Var($\mu \gamma$)	$\gamma = 2.5$	1.25	1.25	$\gamma = 4.5$	2.25	2.25
	$\gamma = 8$	1.40	1.40			
E($\gamma \mu$)	$\mu = -3$	7.5	9	$\mu = -2$	8	10
	$\mu = -1$	8.2	10.58	$\mu = 0$	8	10
	$\mu = 1$	7.2	7.7	$\mu = 2$	6.5	6.6
	$\mu = 3$	5.5	5.5	$\mu = 4$	4.3	4.3
	$\mu = 5$	3.3	3.3	$\mu = 6$	3.3	3.3
	$\mu = 7$	3.2	3.2	$\mu = 8$	3.2	3.2
	$\mu = 9$	3	3	$\mu = 10$	2.9	2.9
	$\mu = 11$	2.9	2.9			
Var($\gamma \mu$)	$\mu = 0$	0.3025	0.3025	$\mu = 3$	0.4225	0.4225
	$\mu = 8$	1.890625	1.890625			

$$m_{12} = 0.1012311, m_{20} = -0.5240978, m_{21} = 0.1412233, m_{22} = -0.01469736.$$

For the (d4T + ddl + IND) treatment, the experts would expect a higher cost and lower effectiveness for treatment with complications. Thus, for an effectiveness of 0, the expected cost would be 10000 euros. Table 3 shows the conditional moments assessed by the expert. The values of the parameters of the prior distribution are:

$$m_{01} = 7.507383, m_{02} = -0.4013869, m_{10} = 67.074594, m_{11} = -1.039000,$$

$$m_{12} = 0.02305853, m_{20} = -0.4792257, m_{21} = 0.1209304, m_{22} = -0.01653587.$$

Figure 1 shows the joint distribution of the prior information on effectiveness and cost for each treatment, together with the joint distribution of the prior incremental effectiveness and cost between treatments. We assume that effectiveness and cost are not the same for treatment without complications as for treatment with complications. There was found to be a bimodal joint distribution for cost and effectiveness. This analysis open a wide range of possibilities for incorporating different prior beliefs distant from the conventional bivariate normal distribution.

The posterior distribution is estimated using Markov Chain Monte Carlo methods. Figure 2 shows the measure used to take decisions. No further comment is made on the results of the analysis because they are not comparable, as the prior information is different for each model.

4 Conclusion

The Bayesian approach to cost–effectiveness analysis has been growing rapidly in popularity as an alternative to the classical approach in the economic evaluation of health technologies. For instance, in the United Kingdom the National Institute for Clinical Excellence (NICE) specifically accepts Bayesian approaches in its guidance to sponsors on making submissions.

The Bayesian approach allows the incorporation of prior information. In a fully Bayesian analysis, the procedures used to elicit expert opinion are an active research issue. The most common prior structure for mean effectiveness and mean cost is the bivariate normal structure.

In this paper we study the use of a more general family of prior distributions for the mean of the effectiveness and cost. In particular, we assume that the conditional density of the mean effectiveness for a given mean cost and the conditional density of the mean cost for a given mean effectiveness are both normal.

The improvement gained over the use of conditionally specified priors is the wide range of prior information that may be elicited. Prior information of more than one source, or different structures of effectiveness and costs depending on whether complications occur, are some cases whereby a conventional bivariate prior distribution is not enough to specify the prior information.

A practical example with real data shows the flexibility of this analysis, incorporating a wide range of possible prior knowledge. Conventional cases, such as the independence case and bivariate prior information, are included as particular cases of this more general analysis.

The posterior distribution is easily simulated using Markov Chain Monte Carlo techniques.

Acknowledgments

Research partially funded by SEJ2006–12685 (MEC, Spain). We thank Xavier Badía for allowing us to use his data archive. All responsibility for the further analysis of the data, of course, is ours alone.

References

- [1] Al. M. & Van Hout, B. (2000). Bayesian approach to economic evaluation of clinical trials: The case of Stenting versus Balloon Angioplasty. *Health Economics*, **9**, 599–609.
- [2] Arnold, B., Castillo, E. & Sarabia, J. (1998). Bayesian analysis for classical distributions using conditionally specified priors. *Sankhya – The Indian Journal of Statistics*, **60**, 228–245.
- [3] Arnold, B., Castillo, E. & Sarabia, J. (2001a). Bayesian inference using conditionally specified priors. *Handbook of Applied Econometrics and Statistical Inference*, Marcel Dekker, New York.
- [4] Arnold, B., Castillo, E. & Sarabia, J. (2001b). Conditionally specified distributions: an introduction. *Statistical Science*, **16**, 249–274.
- [5] Arnold, B., Castillo, E., Sarabia, J. & González-Vega, L. (2000). Multiple modes in densities with normal conditionals. *Statistics and Probability Letters*, **49**, 355–363.
- [6] Brooks, R. & EuroQol group (1996). EuroQol: The current state of play. *Health Policy*, **37**, 53–72.
- [7] Brophy, J. & Joseph, L. (1995). Placing trials in context using Bayesian analysis: Gusto revisited by Reverend Bayes. *Journal of the American Medical Association*, **23**, 871–875.
- [8] Castillo, E. & Galambos, J. (1989). Conditional distributions and the bivariate normal distribution. *Metrika*, **36**, 209–214.
- [9] Chilcott, J., McCabe, C., Tappenden, P., O’Hagan, A., Cooper, N., Abrams, K. & Claxton, K. (2003). Modelling the cost–effectiveness of interferon beta and glatiramer acetate in the management of multiple sclerosis. *British Medical Journal*, **326**, 522–526.
- [10] Frybak, D., Chinnis, J. & Ulvila, J. (2001). Bayesian cost–effectiveness analysis. An example using the GUSTO trial. *International Journal of Technology Assessment in Health Care*, **17**, 83–97.
- [11] Gelman, A. & Meng, X. (1991). A note on bivariate distributions that are conditionally normal. *American Statistics*, **45**, 125–126.
- [12] Heitjan, D. (1997). Bayesian interim analysis of phase II cancer trials. *Statistics in Medicine*, **16**, 1791–1802.
- [13] Heitjan, D., Moskowitz, A. & William, W. (1999). Bayesian estimation of cost–effectiveness ratios from clinical trials. *Health Economics*, **8**, 191–201.

- [14] Jones, D. (1996). *Bayesian approach to the economic evaluation of health care technologies*. In B. Spiker (Ed.), *Quality of Life and Pharmacoeconomics in Clinical Trials* (1189–1196). Philadelphia: Lippincott-Raven.
- [15] Laska, E., Meisner, M. & Siegel, C. (1997). Statistical inference for cost-effectiveness ratios. *Health Economics*, **6**, 229–242.
- [16] Leal, J., Wordsworth, S., Legood, R., & Blair, E. (2007). Eliciting expert opinion for economic models: An applied example. *Value in Health*, **10**, 195–203.
- [17] O’Hagan, A. & Stevens, J. (2001). A framework for cost-effectiveness analysis from clinical trial data. *Health Economics*, **10**, 303–315.
- [18] O’Hagan, A., Stevens, J. & Montmartin, J. (2001). Bayesian cost-effectiveness analysis from clinical trial data. *Statistics in Medicine*, **20**, 733–753.
- [19] Pinto, J., López, C., Badía, X., Corna, A. & Benavides, A. (2000). Análisis coste-efectividad del tratamiento antirretroviral de gran actividad en pacientes infectados por el VIH asintomáticos. *Medicina Clínica*, **114**, 62–67.
- [20] Spiegelhalter, D., Feedman, L. & Parmar, M. (1994). Bayesian approaches to randomized trials (with discussion). *Journal of the Royal Statistical Society, Series A*, **157**, 357–416.
- [21] Spiegelhalter, D., Thomas, A. & Best, N. (2000). WinBUGS version 1.3. User manual, Biostatistics Unit, Cambridge Medical Research Council.
- [22] Stevens, J. & O’Hagan, A. (2002). Incorporating of genuine prior information in cost-effectiveness analysis of clinical trial data. *International Journal of Technology Assessment in Health Care*, **18**, 782–790.
- [23] Stevens, J., O’Hagan, A. & Miller, P. (2003). Case study in Bayesian analysis of a cost-effectiveness trial in the evaluation of health care technologies: Depression. *Pharmaceutical Statistics*, **2**, 51–68.
- [24] Stinnett, A. & Mullahy, J. (1998). Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. *Medical Decision Making*, **18**, S65–S80.
- [25] Tambour, M., Zethraeus, N. & Johannesson, M. (1998). A note on confidence intervals in cost-effectiveness analysis. *International Journal of Technology Assessment in Health Care*, **14**, 467–471.
- [26] Vanness, D. & Kim, W. (2002). Bayesian estimation, simulation and uncertainty analysis: The cost-effectiveness of Ganciclovir Prophylaxis in liver transplantation. *Health Economics*, **11**, 551–566.

- [27] Willan, A. & O'Brien, B. (1996). Confidence intervals for cost-effectiveness ratios: an application of Fielle's theorem. *Health Economics*, **5**, 297–305.