UNIVERSITÀ DEGLI STUDI ROMA TRE
DIPARTIMENTO DI ECONOMIA

A BAYESIAN SEMI-PARAMETRIC APPROACH FOR COST-EFFECTIVENESS ANALYSIS IN HEALTH ECONOMICS

Caterina Conigliani and Andrea Tancredi

Working Paper n° 46, 2005

UNIVERSITÀ DEGLI STUDI ROMA TRE
DIPARTIMENTO DI ECONOMIA

Working Paper n° 46, 2005

- I "Working Papers" del Dipartimento di Economia svolgono la funzione di divulgare tempestivamente, in forma definitiva o provvisoria, i risultati di ricerche scientifiche originali. La loro pubblicazione è soggetta all'approvazione del Comitato Scientifico.

- Per ciascuna pubblicazione vengono soddisfatti gli obblighi previsti dall'art. 1 del D.L.L. 31.8.1945, n. 660 e successive modifiche.

- Copie della presente pubblicazione possono essere richieste alla Redazione.

REDAZIONE:
Dipartimento di Economia
Università degli Studi di Roma Tre
Via Ostiense, 139 - 00154 Roma
Tel. 0039-6-57374003 fax 0039-6-57374093
E-mail: dip_eco@uniroma3.it

UNIVERSITÀ DEGLI STUDI ROMA TRE
DIPARTIMENTO DI ECONOMIA

A BAYESIAN SEMI-PARAMETRIC APPROACH FOR COST-EFFECTIVENESS ANALYSIS IN HEALTH ECONOMICS

Caterina Conigliani and Andrea Tancredi

CONTENTS

Abstract. *We consider the problem of assessing new and existing technologies for their cost-effectiveness in the case where data on both costs and efficacy are available from a clinical trial, and we address it by means of the cost-effectiveness acceptability curve in the simple case where efficacy is measured as a binary outcome. We consider a Bayesian approach, and in recognising that cost data usually exhibit highly skew, heavy-tailed and, possibly multi-modal distributions, we introduce a model for costs composed of a piecewise constant density up to an unknown endpoint, and a generalised Pareto distribution for the remaining tail.*

1 Introduction

The increasing burden on the budgets of health care providers has resulted in considerable interest in assessing new and existing technologies for their clinical effectiveness and cost-effectiveness.

Suppose that two health care technologies, treatment 1 (T_1) and treatment 2 (T_2), are to be compared in a randomised controlled trial; data are direct measurements of efficacy and cost:

$$D = \left\{ x_{ij} = (e_{ij}, c_{ij})^T : \ i = 1, 2; \ j = 1, 2, ..., n_i \right\}$$

where e_{ij} is the efficacy measure of treatment i on patient j and c_{ij} is the cost for patient j under treatment i.

In order to assess if T_1 is more cost-effective than T_2, we need to compare expected efficacies γ_i and expected costs μ_i for each treatment. Let $\Delta_e = \gamma_1 - \gamma_2$ and $\Delta_c = \mu_1 - \mu_2$ be the efficacy and cost differentials. Moreover, let K be a decision-maker's *willingness to pay* coefficient, that is the units of money a decision maker is prepared to pay to obtain one unit of effectiveness.

The primary measure of cost-effectiveness of T_1 relative to T_2 is usually considered to be the *incremental cost-effectiveness ratio*, defined as $\rho = \Delta_c/\Delta_e$. However, as pointed out for instance in O'Hagan *et al.* (2000), cost-effectiveness of T_1 does not simply equate to ρ being less than K. It also depends on the sign of Δ_e, so that it is the sign of the *net monetary benefit* $K\Delta_e - \Delta_c$ that is of interest: T_1 is cost-effective relative to T_2 if $K\Delta_e - \Delta_c > 0$, *i.e.* if in the plane of possible pairs of values of the population mean increments of efficacy and cost, (Δ_e, Δ_c) is below a sloping line of gradient K (see Figure 1).

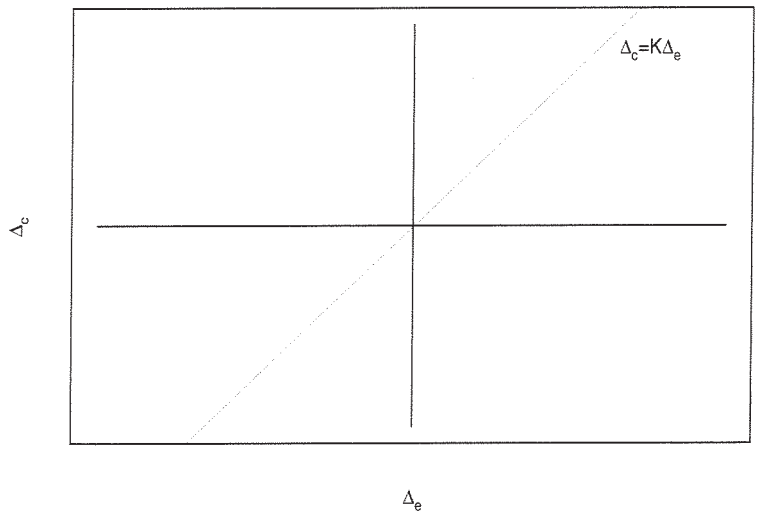

Figure 1: The incremental cost-effectiveness plane

This is usually referred as the *Net Benefit approach* (Stinnett and Mullahy, 1998), and inference about the net monetary benefit is generally presented by means of a Cost-Effectiveness Acceptability Curve (van Hout *et al.*, 1994), that plots the probability $Q(K)$ that the net benefit is positive against the coefficient K, which is rarely unambiguously determined in practice. In this sense, a Bayesian approach is particularly natural, since no such probability exists or has any meaning in frequentist statistics (see O'Hagan *et al.*, 2000, and the references therein), so that $Q(K)$ is the posterior probability given data D

$$Q(K) = P(K\Delta_e - \Delta_c > 0 \mid D).$$

2 The model

2.1 Modelling efficacies

We assume that the clinical outcome of the trial is binary:

$$e_{ij} = \begin{cases} 1 & \text{if } T_i \text{ was effective on patient } j \text{ (positive outcome)} \\ 0 & \text{if } T_i \text{ was not effective on patient } j \text{ (negative outcome)} \end{cases}$$

and let ϕ_i be the probability of positive outcome under T_i.

Then the population mean efficacy under T_i is

$$\gamma_i = \phi_i$$

and the distribution for a single observation x_{ij} under T_i is:

$$g_i\left(x_{ij}\left|\theta_i^0, \theta_i^1, \phi_i\right.\right) = \begin{cases} (1-\phi_i) f_i^0\left(c_{ij}\left|\theta_i^0\right.\right) & \text{if } e_{ij} = 0 \\ \phi_i f_i^1\left(c_{ij}\left|\theta_i^1\right.\right) & \text{if } e_{ij} = 1 \end{cases}$$

where $f_i^0\left(c_{ij}\left|\theta_i^0\right.\right)$ is the distribution for c_{ij} given a negative outcome under T_i and $f_i^1\left(c_{ij}\left|\theta_i^1\right.\right)$ is the distribution for c_{ij} given a positive outcome under T_i.

2.2 Modelling costs

In cost evaluations conceived to have an impact on medical policy, the main interest is the total healthcare cost, so that it is inference on population mean costs that is informative. However, cost data obtained for individual patients in health economic studies typically exhibit highly skew and heavy tailed distributions, and many problems arise with the various approaches currently available for analysing such data.

In fact, as discussed in O'Hagan and Stevens (2002, 2003), non-parametric methods, such as those based on the asymptotic normality of the sample mean or non-parametric bootstrapping, may be inefficient and their justification breaks down in small samples. On the other hand, parametric modelling may lead to more efficient inference, but is dependent on the population distribution matching the model adequately. The main difficulty in this sense, as pointed out for instance in Nixon and Thompson (2004), is that the high skewness and kurtosis usually found in cost data imply that the population mean can be very sensitive to the tail of the distribution, that it is quite complicated to model accurately. One consequence of this is that parametric models that fit the data equally well can produce very different answers; conversely, in some cases models that fit badly can give similar inferences to those that fit well. For these reasons, Nixon and Thompson (2004) recommend that the sensitivity of conclusions to the choice of the model is always investigated, so that model uncertainty becomes a crucial aspect of analysing cost data. Another problem related to the parametric modelling of costs concerns possible transformations of the data; in fact, as discussed in Thompson and Barber (2000) and Briggs and

Gray (1998), mean values and confidence limits may be difficult to interpret on the transformed scales, and back-transformation onto the original scale is not always straightforward.

Here, for each treatment group i and for each efficacy group k ($i = 1, 2;\ k = 0, 1$), we model the bulk of the data and the right tail separately. More specifically, we consider a distribution composed of a piecewise constant density up to an unknown endpoint α, and a generalised Pareto distribution (GPD) for the remaining tail data. The first component of the model, the step function, is very flexible, in the sense that it has the appealing property of catching all the relevant features of the data; if for instance the data exhibit multimodality, the corresponding model will be multimodal. However, the step function will hardly give any weight to values beyond the range of the data; for this reason, we introduce a different model for the right tail of the distribution, the GPD, that is often used in extreme value theory to model tail data (see, for instance Coles, 2001). This mixture model has been applied to environmental data by Tancredi *et al.* (2002) and to cost data arising in the evaluation of health care technologies by Conigliani and Tancredi (2005).

The density function for c_{ij} can then be written as

$$f_i^k(c_{ij}|\theta_i^k) = \begin{cases} (1-\omega)\, h\left(c_{ij}\,|s, p^{(s)}, a^{(s)}, \alpha\right) & 0 < c_{ij} < \alpha \\ \omega\ g\left(c_{ij}\,|\alpha, \sigma, \xi\right) & \alpha \leq c_{ij} < \infty \end{cases} \tag{1}$$

where ω is the probability that an observation c_{ij} is greater than α, $g\left(c_{ij}\,|\alpha, \sigma, \xi\right)$ is the generalized Pareto density with threshold α, scale parameter σ and shape parameter ξ:

$$g\left(c_{ij}\,|\alpha, \sigma, \xi\right) = \frac{1}{\sigma}\left[1 + \frac{\xi\,(x-\alpha)}{\sigma}\right]^{-\frac{1}{\xi}-1}, \tag{2}$$

and $h\left(c_{ij}\,|s, p^{(s)}, a^{(s)}, \alpha\right)$ is a piecewise constant density on $(0, \alpha)$, with unknown upper end point α and unknown number of steps s at positions $a^{(s)} = (a_2, ..., a_s)$, that can be seen as a mixture of s uniform distributions $U_{[a_i, a_{i+1})}$:

$$h\left(c_{ij}\,|s, p^{(s)}, a^{(s)}, \alpha\right) = \sum_{i=1}^{s} p_i U_{[a_i, a_{i+1})}(c_{ij}) \tag{3}$$

with (unknown) weights $p^{(s)} = (p_1, ..., p_s)$ such that $\sum_{i=1}^{s} p_i = 1$.

Note that the parameter vector $\theta_i^k = \left(s, p^{(s)}, a^{(s)}, \alpha, \sigma, \xi, \omega\right)$ varies both with the treatment group i and with the efficacy group k ($i = 1, 2;\ k = 0, 1$). Moreover notice

that for all i and k it is reasonable to assume the constraint $\xi \in [0, 1)$ (see Conigliani and Tancredi, 2005).

It follows that the expected mean cost for treatment i and efficacy k can be written as

$$\mu_i^k = (1-\omega) \sum_{i=1}^{s} p_i \frac{a_{i+1} + a_i}{2} + \omega \left[\alpha + \frac{\sigma}{1-\xi} \right]$$

and the population mean cost under treatment i is

$$\mu_i = (1 - \phi_i)\, \mu_i^0 + \phi_i \mu_i^1.$$

2.3 Prior assumptions and computational issues

First, notice that when approximating the distribution of a tail with the GPD, the scale parameter σ depends on the threshold α (see Coles, 2001), and one should take this into account when introducing a prior distribution for the parameters of the model. However, as pointed out in Tancredi *et al.* (2002) and in Coles and Tawn (1996), this problem can easily be avoided by reparametrizing model (1) in terms of the parameters of the generalized extreme value distribution (GEV) corresponding to the GPD (Coles, 2001). In particular we can write σ and ω as:

$$\sigma = \psi + \xi\,(\alpha - m)$$

and

$$\omega = \frac{1}{n} \left[1 + \frac{\xi\,(\alpha - m)}{\psi} \right]^{-\frac{1}{\xi}}$$

where m and ψ are the location and scale parameters of the GEV, and n is the sample size; this parametrization has the advantage of making the parameters m, ψ, ξ independent of α.

Then, for both treatment groups and for both efficacy groups, we assume that the parameters of the step function are independent of the remaining parameters of the model. Moreover, we express weak prior informations by introducing a uniform prior on $[0, 1]$ for ϕ_1 and ϕ_2, a uniform prior on the range of the data for the threshold α, a uniform prior between 1 and a maximum number of steps $s_{\max}$ for s, and a Dirichlet distribution with all parameters set equal to 1, *i.e.* a uniform prior over the region $\sum_{i=1}^{s} p_i = 1$, for the weights $p^{(s)}$. Then, following Green (2000), as a prior distribution for the step positions $a^{(s)}$ we consider a joint density providing a slight

preference against two step positions occurring too closely in succession:

$$\pi\left(a^{(s)}\,|s,\alpha\right) \propto a_2\left(a_3-a_2\right)...\left(\alpha-a_s\right). \tag{4}$$

Finally, we assume a standard non-informative prior for m and ψ:

$$\pi\left(m,\psi\right)=\pi\left(m\right)\pi\left(\psi\right)\propto\frac{1}{\psi} \tag{5}$$

while we model the shape parameters ξ under the two treatment groups and the two efficacy groups as exchangeable by introducing the hierarchical prior:

$$\begin{array}{c}\log\left(\dfrac{\xi}{1-\xi}\right)\sim N\left(\varepsilon,\eta\right)\\ \varepsilon\sim N\left(0,1\right)\end{array} \tag{6}$$

where ε is estimated using data from all four groups, and η is a known constant. In fact, as pointed out in O'Hagan and Stevens (2003) and in Conigliani and Tancredi (2005), although we might have little prior information, we would not expect the tails of the distributions of costs to be very different between the two treatments. In general the large costs observed under one treatment are indicative of the extreme skewness and kurtosis of costs generally, and their presence suggest that in larger samples one might find comparable skewness and kurtosis also under other treatments.

Note that Bayesian inference for this model is possible using *Markov chain Monte Carlo* (MCMC) methods (see, for instance, Robert and Casella, 1999). In particular, the complete updating of the vector of unknown parameters can be made by using a Gibbs kernel for $p^{(s)}$, the Metropolis-Hastings algorithm for $a^{(s)}$, α, the GEV parameters and ϕ_i $(i=1,2)$, and the reversible jump methodology (Green, 2000) for updating s by splitting one step in two or combining two into one. Details of the algorithm can be found in Tancredi *et al.* (2002).

3 An Example

We present an example using a hypothetical study comparing two treatments in the medical context. The data set is hypothetical but is based on a real example. We assume that a total of 891 patients were randomised in a 2:1 ratio to treatment 1 and treatment 2 respectively, and were followed for one year starting from their date of randomisation. Moreover, we assume that a positive outcome was observed for

Table 1: Healthcare costs: sample descriptive statistics

	Treatment	**1**	**Treatment**	**2**
efficacy group	1	0	1	0
sample size	301	293	105	192
mean	1403	1728	1406	1458
standard deviation	967	1505	1254	1274
median	1157	1411	1111	1244
minimum	26	38	87	20
maximum	9872	13948	9623	13361
skewness $(\overline{\mu}_3/\sigma^3)$	3.7	4.7	4.0	4.9
kurtosis $(\overline{\mu}_4/\sigma^4)$	27	35	24	42

301 patients (out of 594) in the first treatment group and for 105 (out of 297) in the second treatment group. The aim is the analysis of the cost-effectiveness of T_1 compared with T_2.

Table 1 shows sample summaries for the annual healthcare costs. For all treatment groups and efficacy groups, the standard deviations are large, indicating that the data are spread quite far around the mean, and the median cost is smaller than the mean, indicating positively skew data; this fact is confirmed also by the standard skewness statistic $\overline{\mu}_3/\sigma^3$, and by visual inspection of the data in Figure 2 and Figure 3. Finally, the kurtosis statistic indicates that the four distribution of costs are significantly leptokurtic.

We now apply our model to these data in order to assess if T_1 is more cost-effective than T_2. Table 2 shows a 95% posterior credible interval ($PCI_{0.95}$) and a point estimate for some of the parameters of the model (the mean costs μ_1 and μ_2, the mean efficacies ϕ_1 and ϕ_2, and the differentials Δ_c and Δ_e), obtained with 100000 iterations of the simulation algorithm discussed in Section 2, letting $\eta = 0.5^2$ and $s_{max} = 20$. The posterior mean of the mean efficacy is higher for patients treated with T_1, indicating that the first treatment is more effective; in fact, looking at the distribution of the efficacy differential, we obtained $\mathbf{P}(\mathbf{\Delta_e} > \mathbf{0} \mid D) = 0.99$. On the other hand, the posterior mean of the mean cost is higher for patients treated with T_1, so that there is also evidence that the first treatment is more expensive; in fact, looking at the distribution of the cost differential, we obtained $\mathbf{P}(\mathbf{\Delta_c} > \mathbf{0} \mid D) = 0.74$.

Moreover, we computed the Cost-effectiveness acceptability curve, that is plotted

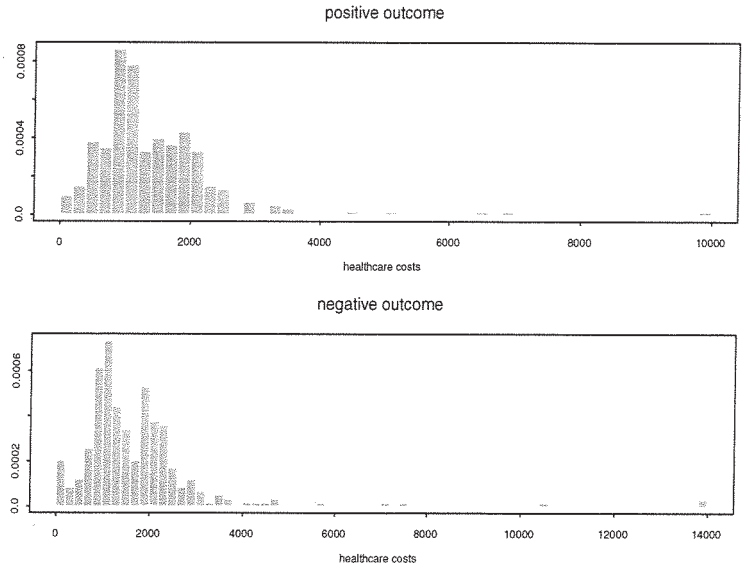

Figure 2: Treatment 1: Healthcare cost data

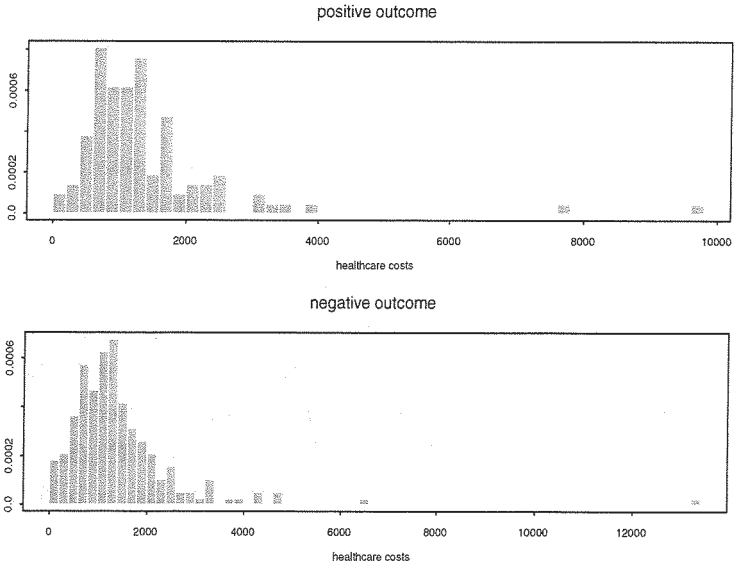

Figure 3: Treatment 2: Healthcare cost data

Table 2: Posterior summaries

			Parameter			
	μ_1	μ_2	Δ_c	ϕ_1	ϕ_2	Δ_e
Mean	1583	1522	60	0.51	0.35	0.15
$\mathbf{PCI_{0.95}}$	1474; 1749	1346; 1829	−224; 272	0.47; 0.55	0.30; 0.41	0.08; 0.22

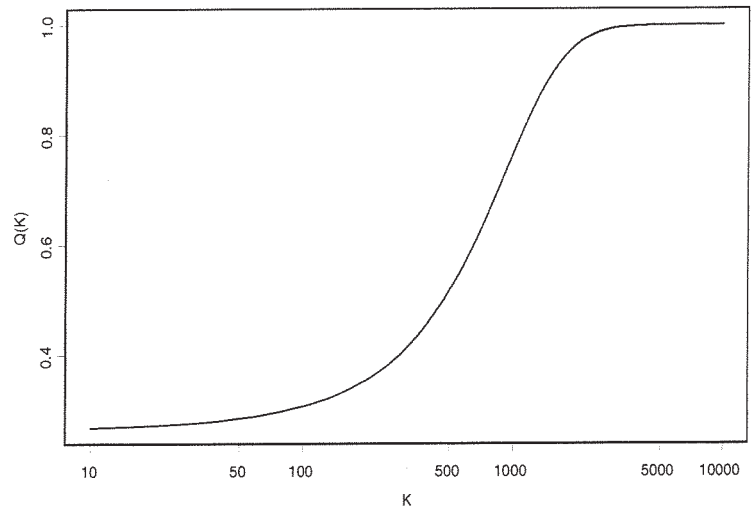

Figure 4: Cost-Effectiveness Acceptability Curve

in Figure 4. We see that $Q(K)$ is 0.26 at $K = 0$, corresponding to the probability that T_1 is cheaper, and tends to 1 as $K \to \infty$, corresponding to the probability that T_1 is more effective. If the health provider were willing to pay more than 1090 to obtain one unit of effectiveness, then there is a posterior probability of at least 80 per cent that T_1 is more cost-effective than T_2. However, for smaller values of K the case in favour of T_1 is much less clear.

4 Discussion

We have considered the problem of assessing new and existing technologies for their cost-effectiveness in the case where data on both costs and efficacy are available from

a clinical trial. We have addressed this problem by means of the cost-effectiveness acceptability curve in the simple case where efficacy is measured as a binary outcome, introducing a semi-parametric distribution for modelling cost data, that combines the semi-parametric approach to density estimation based on mixture models and the semi-parametric approach to tail estimation based on extreme value theory.

However, the clinical outcome of trials usually is not binary; in the example of Section 3, for instance, the primary clinical endpoint was symptom-free days. It follows that more accurate results in terms of cost-effectiveness can be obtained by modelling the data more carefully with an appropriate two dimensional distribution. In particular, by embedding model (1) in a two dimensional distribution, our approach can be used also to model jointly costs and efficacy measures.

References

1. Briggs A, Gray A. The distribution of health care costs and their statistical analysis for economic evaluation. *J. Health Serv. Res. Policy.* 1998; 3, 233-245.
2. Coles SG. *An Introduction to Statistical Modeling of Extreme Values.* Springer-Verlag, London, 2001.
3. Coles SG, Tawn JA. A Bayesian analysis of extreme rainfall data. *Appl. Stat.* 1996; 45, 463-478.
4. Conigliani C, Tancredi A. Semi-parametric modelling for costs of health care technologies, *Statist.Med.*, to appear, 2005.
5. Green PJ. A Primer on Markov Chain Monte Carlo. In *Complex Stochastic Systems*, edited by O.E. Barndorff-Nielsen, D.R. Cox, C. Kluppelberg, Chapman & Hall. 2000.
6. Nixon RM, Thompson SG. Parametric modelling of cost data in medical studies. *Statist.Med.* 2004; 23: 1311-1331.
7. O'Hagan A, Stevens JW. Bayesian methods for design and analysis of cost-effectiveness trials in the evaluation of health care technologies. *Statistical Methods in Medical Research* 2002; 11,469-490.
8. O'Hagan A, Stevens JW. Assessing and comparing costs: how robust are the bootstrap and methods based on asymptotic normality? *Health Economics* 2003; 12, 33-49.

9. O'Hagan A, Stevens JW., Montmartin J. Inference for the Cost-Effectiveness Acceptability Curve and Cost-Effectiveness Ratio, *Pharmacoeconomics* 2000; 17, 339-349.

10. Robert CP, Casella G. *Monte Carlo Statistical Methods* Springer-Verlag, New York. 1999.

11. Stinnett AA., Mulahy J. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis, *Medical Decision Making*, 1998; 18: 68-80.

12. Tancredi A, Anderson CW, O'Hagan A. A Bayesian model for threshold uncertainty in extreme value theory. *Working paper 2002.12, Dipartimento di Scienze Statistiche, Università di Padova. Available at http://3w.eco.uniroma1.it/utenti/tancredi/* . Submitted

13. Thompson SG, Barber JA. How should cost data in pragmatic randomised trials be analysed? *Br. Med. J.* 2000; 329, 1197-1200.

14. van Hout BA, Al MJ, Gordon GS, et al. Costs, effects and C/E ratios alonside a clinical trial *Health Economics*, 1994; 3:309-19.

Finito di stampare nel mese di luglio 2005, presso *Tipolitografia artigiana* **Colitti Armando** snc
00154 Roma • Via Giuseppe Libetta 15 a • Tel. 065745311/065740258
e-mail tcolitti@tin.it • www.colitti.it


GIURII