Statistical issues in Bayesian cost–effectiveness analysis

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

E. Moreno, Dpt. of Statistics, University of Granada, Spain

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

Dpt. of Quantitative Methods, University of Las Palmas de Gran Canaria, Spain

emoreno@ugr.es,

{francisco.vazquezpolo or miguel.negrin or maria.martel}@ulpgc.es

Abstract

Cost–effectiveness analysis of medical treatments search for choosing an "optimal" treatment among a set of $k \geq 2$ alternative treatments $T_1, ..., T_k$ for a given disease. It is imposed that the cost and the effectiveness of the treatments are taken into account in the selection procedure.

We focus the problem as a Bayesian statistical decision problem, present their elements and illustrate the procedure. Further, we discuss some difficulties arising in cost–effectiveness analysis when heterogeneity is present in the cost and effectiveness data. Heterogenous data implies in cost–effectiveness analysis the need of considering special statistical techniques such as Bayesian meta–analysis and Bayesian probabilistic clustering.

Keywords and phrases. Cost and effectiveness of a treatment, predictive reward distribution of a treatment, optimal treatment, utility function.

1 Introduction

Health Economics is an area of the field of Economics with an intensive recent development. The major concerns of researchers in this area is the comparison between medical treatments based on their effectiveness and cost. It is accepted that health resources are limited and effectiveness comes at a price. As control over health expenditure has increased over the last thirty years, the term cost-effectiveness (CEA) has gained in popularity.

This increasing focus on CEA of new or existing treatments has been led by the development of health technology assessment (HTA) agencies, such as the National Institute for Health and Care Excellence (NICE) in the UK, which seeks to provide guidelines for Health care providers and decision makers about which treatments should be covered in a context of scarce economic resources.

In Europe, since 2008, the European Medicines Agency has been working closely with Health Technology Assessment (HTAs) Bodies in different Member States, as well as with the European Network for Health Technology Assessment (EUnetHTA), with the objective of generating relevant data for regulators, HTA bodies and other interested parties.

In the United States, the federal government has provided financial support for health technology assessment since the early 70s. The US Office of Technology Assessment (OTA), the Medicare Coverage Division with the Center for Medicare and Medicaid Services (CMS), and the Agency for Healthcare Research and Quality (AHRQ) are some federal institutions that undertake or fund cost or cost-effectiveness analyses of medical technologies and interventions (see (1) and (2)).

In other countries, such as Australia or Canada, it is regulated that pharmaceutical companies should submit their products to CEA (3).

All the research efforts on cost–effectiveness analysis are spread out on topics that range from the formal definition and measurement of effectiveness and cost of a medical treatment to the development of tools for treatment comparison. The statistical decision theory plays an central role for understanding CEA and in this paper we do briefly summarize some achievements and statistical difficulties in this area.

The rest of the paper is organized as follows. In Section 2 we consider the evolution of the statistical tool for cost effectiveness analysis starting from the direct consideration of the random variables cost c and effectiveness e of a treatment, to the more sophisticated notion of net benefit. In Section 3 we introduce the cost–effectiveness analysis as a decision problem, identify the reward of a treatment, introduce two utility functions and the notion of optimal treatment. Section 4 describes the statistical problems that arise when the samples are not homogenous.

2 Statistical tools for cost-effectiveness analysis

For a time the incremental cost-effectiveness ratio (ICER) was the basic tool for cost-effectiveness analysis (4). Let $\Delta c = Ec_1 - Ec_2$ and $\Delta e = Ee_1 - Ee_2$ be the difference of the expectation of cost and effectiveness of two given treatments T_1 and T_2. Then, the ICER for the treatments is defined as the ratio

$$ICER_{12} = \frac{\Delta c}{\Delta e},$$

whose meaning is the increment of cost per unit of increment of effectiveness of treat-

ments T_1 and T_2. Therefore, it is assumed that cost and effectiveness (c, e) are random variables with a partially unknown distribution $P(c, e)$.

In a cartesian plane with axes $(\Delta e, \Delta c)$ quadrant I (QI) correspondons to $\Delta e \geq 0, \Delta c \geq 0$, quadrant II (QII) to $\Delta e \leq 0, \Delta c \geq 0$, quandrant III (QIII) to $\Delta e \leq 0, \Delta c \leq 0$ and quandrant IV (QIV) to $\Delta e \geq 0, \Delta c \leq 0$. For $ICER_{12}$ in QI treatment T_1 is more costly and more effective than T_2, in QII T_1 is more costly and less effective than T_2, in QIII T_1 is less costly and less effective than T_2, and in QIV T_1 is less costly and more effective than T_2. It is clear that when $ICER_{12}$ is in QII T_2 is preferred to T_1 and if it in QIV then T_1 is preferred to T_2. However, when $ICER_{12}$ is in either in QI or QIII the decision is not so evident. In those cases a subjective input on the cost per unit of increment of effectiveness is necessary.

When the distribution $P(c, e)$ is not completely known $ICER_{12}$ has to be estimated from a sample of cost and effectiveness of patients under treatment T_1 and T_2. Let $\mathbf{c}_i = (c_{i1}, ..., c_{in_i})$ and $\mathbf{e}_i = (c_{i1}, ..., c_{in_i})$ for $i = 1, 2$, be such a samples. An estimation of $ICER_{12}$ that can be very inaccurate is given by

$$\hat{C}r_{12} = \frac{\bar{c}_1 - \bar{c}_2}{\bar{e}_1 - \bar{e}_2},$$

where $\bar{c}_i = \sum_{j=1}^{n_i} c_{ij}/n_i$ and $\bar{e}_i = \sum_{j=1}^{n_i} e_{ij}/n_i$ are the samples means. Suggestions and criticisms on how to measure the uncertainty on the $ICER_{12}$ estimation have been given by many authors. For instance, when the distribution is completely unknown a bootstrap methodology is advocated by Chaudhary and Stearns (5) and Briggs et al. (6). For further discussion on interpretive problems of the ICER see (7) or (8).

An extension of the $ICER$ for comparing two treatments is the incremental net benefit (INB_{21}) introduced by Stinnett and Mullahy (9). This is defined as

$$INB_{12} = R_{12}\Delta e - \Delta c,$$

where R_{12} is the monetary value assigned to the unit of increment of effectiveness of treatments T_1 and T_2. There are obvious relationships between the $ICER_{12}$ and the INB_{12} that we do not discuss here. An estimation of the INB_{12} that can be very inaccurate is given by

$$\hat{I}_{12} = R_{12}(\bar{e}_1 - \bar{e}_2) - (\bar{c}_1 - \bar{c}_2).$$

The sample variance is a mensure on the uncertainty of the estimator $\hat{I}_{12}$ that is,

$$Var(\hat{I}_{12}|R_{12}) = \sum_{i=1}^{2} \frac{R_{12}^2 s_{e_i}^2 + s_{c_i}^2 - 2R_{12}s_{e_i}^2 s_{c_i}^2 r_i}{n_i},$$

where $s_{e_i}^2$, $s_{c_i}^2$ are the sample variances of the effectiveness and cost, and r_i the estimator of the linear correlation.

A third tool in cost–effectiveness analysis is the cost–effectiveness acceptability curve (CEAC) introduced by Van Hout et al. (10). This is a sampling evaluation of $\hat{I}_{21}$, that is, this notion is defined as the function

$$\varphi(R_{12}) = \Pr(\hat{I}_{12} \geq 0 | R_{12})$$

for $R_{12} \geq 0$. We note that for a given R_{12},

$$\varphi(R_{12}) = \int_C d\bar{e}_1 d\bar{e}_2 d\bar{c}_2 d\bar{c}_1$$

where C is the set given by $C = \Big\{(\bar{e}_1, \bar{e}_2, \bar{c}_1, \bar{c}_2) : R_{21}(\bar{e}_2 - \bar{e}_1) - (\bar{c}_2 - \bar{c}_1) \geq 0\Big\}$.

The interpretation of the curve $\varphi(R_{12})$ for a given $R_{12} \geq 0$ is the sampling probability of the event C as the sample means $\bar{e}_1, \bar{e}_2, \bar{c}_1, \bar{c}_2$ vary in their sampling spaces. In the literature this curve is utilized for choosing an "optimal" treatment: T_1 is optimal if $\varphi(R_{12}) \geq 1/2$. We note that this implies that treatment T_1 is chosen regardless the data $\mathbf{c}_i = (c_{i1}, ..., c_{in_i})$ and $\mathbf{e}_i = (c_{i1}, ..., c_{in_i})$ for $i = 1, 2$ we observed. This suggests that CEAC is not an appropriate tool for choosing optimal treatment.

The INB is the most interesting tool although it is restricted to the case of comparing two treatments. For comparing more than two treatments a more general tool is needed. This extension can be formulated using the notion of *net benefit* z of a treatment T. This was introduced in Moreno et al. (11) and can be considered as an extension of the INB_{12}. For a given $R > 0$, the net benefit z is a random variable defined by

$$z = R \times e - c,$$

where R means the quantity the health provider is willing to pay for the unit of effectiveness. This way, for a given set of alternative treatments $T_1, ..., T_k$, $k \geq 2$, and $R > 0$, we have the net benefits $z_1, ..., z_k$ and treatment comparisons is just the comparison of the distributions of the net benefit of the treatments conditional on R. To do that we need the use of a more sophisticated decision theory methodology that we briefly describe in the next Section.

3 Cost–effectiveness analysis as a decision problem

A general theory for CEA follows by focusing this problem as a decision problem. This is the aim of the book by Moreno, Vázquez–Polo and Negrín (12).

Let us assume that for a given disease we have $k \geq 2$ alternative treatments $T_1, ..., T_k$, and the problem is that of choosing an optimal treatment based on their random cost and effectiveness (c, e). The element of this decision problem are i) a finite decision space $D = \{d_1, ..., d_k\}$, where d_j is the decision of choosing treatment T_j, ii) the reward of each decision which is given by the probability distribution of (c, e), that is,

$\{P(c,e|\theta_1),...,P(c,e|\theta_k\}$, where θ_j is a parameter tight to treatment T_j for $j=1,...,k$, and iii) a utility function $U(c,e)$, the utility we obtain for (c,e).

We note that $z(R)$ can be considered a utility function, that is $U_1(c,e|R)=z$. Under this utility function the utility of the reward $P(c,e|\theta_j)$ is

$$U_1(P(c,e|\theta_i)|R)=R\int\int e\ P(c,e|\theta_i)dcde-\int\int c\ P(c,e|\theta_i)\ dc\ de.$$

This utility depends on θ_i and hence the utilities of the rewards $\{P(c,e|\theta_1),...,P(c,e|\theta_k\}$ cannot be compared. Thus, we need to eliminate this parameter from the distribution $P(c,e|\theta_i)$. The Bayesian way requieres two steps:

1.- We first complete the sampling model $P(c,e|\theta_i)$ to the Bayesian model

$$M_i:\{P(c,e|\theta_i),\pi(\theta_i)\},$$

where $\pi(\theta_i)$ is a prior distribution for the parameter θ_i. This prior distribution may contain subjective prior information on θ_i. If prior information on θ_i is not available an objective prior, as the reference prior (13), can be utilized. Then, for model M_i and sample $\mathbf{c}_i=(c_{i1},...,c_{in_i})$ and $\mathbf{e}_i=(c_{i1},...,c_{in_i})$ the updated posterior distribution of θ_i is given by

$$\pi(\theta_i|\mathbf{c}_i,\mathbf{e}_i)=\frac{(\prod_{i=1}^{n_i}P(c_{ij},e_{ij}|\theta_i))\,\pi(\theta_i)}{\int(\prod_{i=1}^{n_i}P(c_{ij},e_{ij}|\theta_i))\,\pi(\theta_i)d\theta_i}.$$

2.- We compute the updated reward distribution of (c,e) for treatment T_i, which is obtained as

$$P(c,e|\mathbf{c}_i,\mathbf{e}_i)=\int P(c,e|\theta_i)\pi(\theta_i|\mathbf{c}_i,\mathbf{e}_i)d\theta_i.$$

Then, the utility of the reward $P(c,e|\mathbf{c}_i,\mathbf{e}_i)$ is given by

$$U_1(z_i|R)=R\int\int e\ P(c,e|\mathbf{c}_i,\mathbf{e}_i)dcde-\int\int c\ P(c,e|\mathbf{c}_i,\mathbf{e}_i)\ dc\ de,$$

that is, it is a linear function of the expected cost and effectiveness.

For a given R, the optimal treatment is T_j if

$$U_1(P(c,e|\mathbf{c}_j,\mathbf{e}_j)|R)=\max_{i=1,...,k}U_1(P(c,e|\mathbf{c}_i,\mathbf{e}_i)|R).$$

Example Let us consider two treatments T_1 and T_2 with normal reward distributions given by

$$P(c,e|\theta_i)=N(c|\mu_{ci},\sigma_{ci}^2)N(e|\mu_{ei},\sigma_{ei}^2),$$

where $(\mu_{ci},\sigma_{ci}^2,\mu_{ei},\sigma_{ei}^2)$, $i=1,2$, are unknown parameters. For the samples $\mathbf{c}_i=(c_{i1},...,c_{in_i})$ and $\mathbf{e}_i=(c_{i1},...,c_{in_i})$ and priors

$$\pi(\mu_{ci},\sigma_{ci})\propto\frac{1}{\sigma_{ci}},\ \pi(\mu_{ei},\sigma_{ei})\propto\frac{1}{\sigma_{ci}},$$

the posterior distributions are given by

$$\pi(\mu_{ci},\sigma_{ci}|\bar{c}_i,s_i^2)=N\left(\mu_{ci}|\bar{c}_i,\frac{\sigma_{ci}^2}{n_i}\right)\frac{(n_is_{ci}^2)^{(n-1)/2}}{2^{(n-3)/2}\Gamma\left(\frac{n_i-1}{2}\right)}\frac{1}{\sigma_{ci}^{n_i}}\exp\left\{-\frac{n_is_{ci}^2}{2\sigma_{ci}^2}\right\}$$

and

$$\pi(\mu_{ei},\sigma_{ei}|\bar{e}_i,s_i^2)=N(\mu_{ei}|\bar{e}_i,\frac{\sigma_{ei}^2}{n_i})\frac{(n_is_{ei}^2)^{(n-1)/2}}{2^{(n-3)/2}\Gamma\left(\frac{n_i-1}{2}\right)}\frac{1}{\sigma_{ei}^{n_i}}\exp\left\{-\frac{n_is_{ei}^2}{2\sigma_{ei}^2}\right\},$$

where $\bar{c}_i=\frac{1}{n_i}\sum_{j=1}^{n_i}c_{ij}$, $s_{ci}^2=\frac{1}{n_i}\sum_{j=1}^{n_i}(c_{ij}-\bar{c}_i)^2$, $\bar{e}_i=\frac{1}{n_i}\sum_{j=1}^{n_i}c_{ij}$, and $s_{ei}^2=\frac{1}{n_i}\sum_{j=1}^{n_i}(e_{ij}-\bar{e}_i)^2$.

Then, the predictive distribution of the cost c of treatnent T_i is given by the generalized Student t distribution

$$P(c|\mathbf{c}_i)=K(n_i,s_{ci}^2)\left(n_is_{ci}^2+\frac{n_i}{n_i+1}(c-\bar{c}_i)^2\right)^{-n_i/2},$$

where

$$K(n_i,s_{ci}^2)=\frac{\Gamma\left(\frac{n_i}{2}\right)}{\Gamma\left(\frac{n_i-1}{2}\right)\Gamma\left(\frac{1}{2}\right)}\frac{n_i^{1/2}}{(n_i+1)^{1/2}}(n_is_{ci}^2)^{(n_i-1)/2},$$

and a similar expression for $P(e|\mathbf{e}_i)$ replacing in this expression s_{ci}^2 with s_{ei}^2 and $\bar{c}_i$ by $\bar{e}_i$.

It can be seen that

$$U_1(P(c|\mathbf{c}_i)|R)=\int\int c\;P(c,e|\mathbf{c}_i,\mathbf{e}_i)dcde=\bar{c}_i,$$

and analogously, $U_1(P(e|\mathbf{e}_i)|R)=\bar{e}_i$. Then, for a given R the optimal treatment is T_1 if the inequality

$$R\,\bar{e}_1-\bar{c}_1\geq R\,\bar{e}_2-\bar{c}_2,$$

holds, and T_2 otherwise.

Non linear utility functions have been considered in the literature. Let Z_1 and Z_2 be the random net benefit of treatments T_1 and T_2 with rewards $P(z_1|R)$ and $P(z_2|R)$. A nonlinear utility function $U_2(z|R)$ is given by

$$U_2(z_1|R)=\Pr(Z_2\leq z_1|R),\;U_2(z_2|R)=\Pr(Z_1\leq z_2|R).$$

The optimality criterium is now that treatment T_1 is optimal for a given R if the inequality

$$\Pr(Z_1\geq Z_2|R)\geq\Pr(Z_2\geq Z_1|R)$$

holds, and T_2 otherwise. This nonlinear utility function is explored in Chapter 4 in (12).

We note that the optimal treatment for the utility function $U_1(z|R)$ do not necessarily coincide with the optimal treatment for $U_2(z|R)$ as the following simple example shows.

Example Let T_1 and T_2 be two treatments with the same deterministic cost $c = 0$, and effectiveness given by a discrete variable with values 0, 1 and 2, as a health indicator of bad, good and excellent status. The distribution of the effectiveness of treatment T_1 is given by

$$P_1(e_1) = \begin{cases} 0.1, & \text{if} \quad e_1 = 0, \\ 0.5, & \text{if} \quad e_1 = 1, \\ 0.4, & \text{if} \quad e_1 = 2, \end{cases}$$

and the distribution of the effectiveness of treatment T_2 by

$$P_2(e_2) = \begin{cases} 0.3, & \text{if} \quad e_1 = 0, \\ 0.1, & \text{if} \quad e_1 = 1, \\ 0.6, & \text{if} \quad e_1 = 2. \end{cases}$$

The rewards of treatments T_1 and T_2 are certainly different although for the utility function $U_1(z|R)$ the utility of $P_1(z_1|R)$ and $P_2(z_2|R)$ is the same, that is, the expectations of z_1 and z_2 are $E_{P_1}(z_1|R) = E_{P_2}(z_2|R) = 0.13R$. This implies that T_1 and T_2 are equivalent treatments for any $R \geq 0$.

However, for the utility function $U_2(z_i|R)$, the utility of $P_1(z_1|R)$ is

$$\Pr(Z_1 \geq Z_2|R) = \Pr(e_1 \geq e_2) = 0.63,$$

and the utility of $P_2(z_2|R)$ is $\Pr(Z_2 \geq Z_1|R) = \Pr(e_2 \geq e_1) = 0.69$.

Therefore, under the utility function $U_2(z_i|R)$ treatment T_2 is preferred to T_1 for any $R \geq 0$.

4 The between sample heterogeneity in CEA

A difficulty in CEA comes from the fact that the samples of cost and effectiveness $\mathbf{c}_i$ and $\mathbf{e}_i$ often come from h different Health care centers, so that they are an aggregate of samples. That is $\mathbf{c}_i = \cup_{j=1}^{h} \mathbf{c}_{ij}$ and $\mathbf{e}_i = \cup_{j=1}^{h} \mathbf{e}_{ij}$, where $\mathbf{c}_{ij} = (c_{i1}, ..., c_{in_{ij}})$ and $\mathbf{e}_{ij} = (e_{i1}, ..., e_{in_{ij}})$ are samples from hospital j. The distribution of $(\mathbf{c}_{ij}, \mathbf{e}_{ij})$ might change as j changes, and hence we might have h different sampling distributions. Therefore, the heterogeneity adds uncertainty on the model for c and e, and a statistical procedure to account for this model uncertainty is called for.

Let $\{P(\mathbf{c}_{ij}, \mathbf{e}_{ij}|\theta_j),\ j = 1, ..., h\}$ be the sampling distributions conditional on the centers, and $\{P(c, e|\mathbf{c}_{ij}, \mathbf{e}_{ij}),\ j = 1, ..., h\}$ the predictive distributions of the centers. The quantity of interest is the predictive distribution of (c, e) of the treatments and hence for each treatment T_i the distributions $\{P(c, e|\mathbf{c}_{ij}, \mathbf{e}_{ij}),\ j = 1, ..., h\}$ have to be pooled. The statistical procedure for pooling these distribution is known as meta–analysis.

Thus, we strongly recommend the use of meta–analysis in CEA for heterogenous data.

On the other hand, it might be that some of the models $\{P(\mathbf{c}_{ij}, \mathbf{e}_{ij}|\theta_j),\ j = 1, ..., h\}$ have the same parameter. Thus, the point is to reduce the number of models by clustering those that have the same parameter. This is known as probabilistic clustering. A Bayesian approach to clustering the samples with the same distribution is based on *product partition models* introduced by Hartigan (14) and further explored in (15), (16) and (17).

Misleading meta–inferences can be obtained when clustering is ignored, as illustrated in (18). Therefore, clustering the samples before to carried out a CEA is a good practice.

Acknowledgement. Research partially funded by grant ECO2017–85577–P.

References

Luce, B.R. and Brown, R.E., The use of technology assessment by hospitals, health maintenance organizations, and third party payers in the United States, International Journal of Technology Assessment in Health Care, 11 (1), 79–92 (1995)

Sullivan, S.D. and Watkins, J. and Sweet, B. and Ramsey, S.D., Health Technology Assessment in Health–Care Decisions in the United States, Value in Health, 12 (2), S39-S44 (2009)

Hayley, D., The history of health technology assessment in Australia, International Journal of Technology Assessment in Health Care, 25 (1), 61–67 (2009)

O'Brien, B.J. and Gertsen, K. and Willan, A.R. and Faulkner, A., Is there a kink in consumers' threshold value for cost-effectiveness in health care?, Health Economics, 11 (2), 175–180 (2002)

Chaudhary, M.A. and Stearns, S.C., Estimating confidence intervals for cost–effectiveness ratios: an example from a randomized trial, Statistics in Medicine, 15 (13), 1447–1458 (1996)

Briggs, A.H. and Wonderling, D.E. and Mooney, C.Z., Pulling cost–effectiveness analysis up by its bootstrap: A non–parametric approach to confidence interval estimation, Health Economics, 6 (4), 327–340 (1997)

Heitjan, D.F. and Moskowitz, A.J. and Whang, W., Bayesian estimation of cost–effectiveness ratios from clinical trials, Health Economics, 8 (3), 191–201 (1999)

O'Hagan, A. and Stevens, J.W. and Montmartin, J., Inference for the cost–effectiveness acceptability curve and cost–effectiveness ratio, Pharmacoeconomics, 17 (4), 339–349 (2000)

Stinnett, A.A. and Mullahy, J., Net health benefits: A new framework for the analysis of uncertainty in cost–effectiveness analysis, Medical Decision Making, 18 (2), S68–S80 (1998)

Van Hout, B.A. and Al, M.J. and Gordon, G.S. and Rutten, F.F., Costs, effects and C/E–ratios alongside a clinical trial, Health Economics, 3 (5), 309–319 (1994)

Moreno, E. and Girón, F.J. and Vázquez–Polo, F.J. and Negrín, M.A., Optimal health-care decisions: comparing medical treatments on a cost–effectiveness basis, European Journal of Operational Research, 204 (1), 180–187 (2010)

Moreno, E. Vázquez–Polo, F.J. and Negrín, M.A.: Cost–effectiveness analysis of medical treatments: a statistical decision theory approach. Chapman & Hall, London, in press (2019)

Berger, J.O. and Bernardo, J.M. and Sun, D., The formal definition of reference priors, Annals of Statistics, 37 (2), 905–938 (2009)

Hartigan, J. A., Partition Models, Communications in Statistics - Theory and Methods, 19, 2745–2756 (1990).

Barry, D. and Hartigan, J.A., Product partition models for change point problems, Annals of Statistics, 20, 260–279 (1992)

Moreno, E., Girón, F.J., Martínez, M.L., Vázquez–Polo, F.J. and Negrín, M.A., Optimal treatments in cost–effectiveness analysis in the precense of covariates: Improving patient subgroup definition, European Journal of Operational Research, 226, 173–182 (2013)

Casella, G., Moreno. E., and Girón, F.J., Cluster Analysis, Model Selection, and Prior Distributions on Models, Bayesian Analysis, 9 (3), 613–658 (2014)

Moreno, E. Vázquez–Polo, F.J. and Negrín, M.A., Bayesian meta–analysis: The role of the between sample heterogeneity, Statistical Methods for Medical Research, in press (2017)