IMA Journal of Management Mathematics (2005) **16**, 239–254
doi:10.1093/imaman/dpi018
Advance Access publication on April 29, 2005

A methodology for measuring the relative effectiveness of healthcare services

GIORGIO VITTADINI†
Department of Statistics, University of Milano-Bicocca, 20126 Milano, Italy

AND

SIMONA C. MINOTTI‡
Department of Economics and Social Sciences, Catholic University of Piacenza, 29100 Piacenza, Italy

In this paper we propose a methodology for measuring the 'relative effectiveness' of healthcare services (i.e. the effect of hospital care on patients) under general conditions, in which: α) a healthcare outcome underlies qualitative and quantitative observable indicators; β) we are interested in studying the simultaneous dependency of multiple outcomes on covariates (where the outcomes can also be correlated to each other); γ) the relative effectiveness is adjusted for hospital-specific covariates; δ) we hypothesise a general distribution for random disturbances and the random parameters of relative effectiveness. For this topic, a generalisation of the SURE (seemingly unrelated regression equations) multilevel model is proposed. The solutions are obtained by means of Bayesian inference methods. Since there is currently no software available to estimate this model, an SAS procedure based on Markov Chain Monte Carlo methods has been developed by the authors, in line with Goldstein & Spiegelhalter (1996, *J. R. Stat. Soc. Ser. A*, **159**, 385–443), Spiegelhalter *et al.* (1996, *Bayesian Using Gibbs Sampling Manual*. Cambridge: MRC Biostatistic Unit, Institute of Public Health) and Albert & Chib (1997, *J. Am. Stat. Assoc.*, **92**, 916–925). In addition, a new theoretical result regarding the joint posterior distribution for the parameters is provided. The model proposed has been implemented for an effectiveness study of a selection of Lombard hospitals.

Keywords: relative effectiveness; multilevel model; seemingly unrelated regression equations; SURE multilevel model; Markov Chain Monte Carlo; Gibbs sampling.

1. Introduction

Performance measurement for healthcare services, i.e. hospitals, is becoming increasingly important for the improvement of healthcare quality.

In this paper we refer to the concept of 'effectiveness', i.e. the effect of hospital care on patients. We are particularly interested in measuring 'relative effectiveness' in order to compare different healthcare institutions in terms of 'healthcare outcomes'.

A healthcare outcome is definable as the "technical result of a diagnostic procedure or specific treatment episode" (Opit, 1993).

Healthcare outcomes, however, are influenced by covariates concerning the 'case-mix' of the patients, definable as the variability of their clinical and socio-demographic aspects. In addition, healthcare outcomes are related to the organisational capacity, resources, facilities and other characteristics of hospitals (Zaslavsky, 2001). Therefore, to allow comparisons between healthcare institutions, what

†Email: giorgio.vittadini@unimib.it
‡Email: simona.minotti@unicatt.it

is of major interest is not only the standardisation of healthcare outcomes but also the influence of the individual and institutional covariates. Relative effectiveness needs to be adjusted for patient-specific and hospital-specific variables, as underlined by Goldstein & Spiegelhalter (1996): "When standard measures are available across organisations, they should be risk-adjusted to account for differences in environmental factors influencing outcomes".

In the nineties, numerous authors (Thomas *et al*., 1994; Normand *et al*., 1995; Morris & Christiansen, 1996; Goldstein & Spiegelhalter, 1996; Rice & Leyland, 1996; Leyland & Boddy, 1998; Marshall & Spiegelhalter, 2001) proposed studying 'relative effectiveness' by means of the multilevel model (Hox, 1995). "The multilevel model overcomes small sample problems by appropriately pooling information across institutions, introducing some bias or *shrinkage*, and providing a statistical framework that allows one to quantify and explain variability in outcomes through the investigation of institutional level covariates" (Marshall & Spiegelhalter, 2001, p. 128).

The multilevel model specified for the j-th outcome is:

$$\boldsymbol{y}_j = \boldsymbol{u}_j \boldsymbol{Z}_j + \boldsymbol{\beta}_j \boldsymbol{X}_j + \boldsymbol{e}_j, \tag{1}$$

where $\boldsymbol{y}_j$ is a row vector $(1 \times n)$ which contains the values of the j-th outcome regarding the n patients hospitalised in participating q hospitals ($n = \sum_{\upsilon=1}^{q} n_\upsilon$, where n_υ is the number of patients in the υ-th hospital); $\boldsymbol{u}_j$ is a row vector $(1 \times q)$ containing random coefficients interpretable as the effectiveness of hospitals (with respect to outcome $\boldsymbol{y}_j$) adjusted for patient characteristics; $\boldsymbol{Z}_j$ is a 'design matrix' $(q \times n)$; $\boldsymbol{\beta}_j$ is a row vector $(1 \times t_j)$ containing fixed coefficients for patient covariates; $\boldsymbol{X}_j$ is a non-stochastic matrix $(t_j \times n)$ of covariates referring to hospitalised patients (these covariates also include the initial state of health); $\boldsymbol{e}_j$ is a row vector $(1 \times n)$ containing random disturbances associated with the patients.

The authors cited above have considered the multilevel model under the following restrictions:

a) the consideration of one outcome at a time as the response variable;
b) binomial distribution of the outcome;
c) no consideration of hospital-specific characteristics as possible covariates;
d) assumption of multivariate normal distribution for random disturbances and the random parameters of effectiveness.

However, the restrictions indicated above often reduce the applicability of effectiveness studies when:

α) a healthcare outcome, which can be expressed in statistical terms as a latent construct, underlies qualitative and quantitative observable indicators;
β) we are interested in studying the simultaneous dependency of multiple outcomes on covariates, where the outcomes can also be correlated to each other (e.g. the correlation between mental and physical states of health in quality of life studies);
γ) relative effectiveness needs to be adjusted for hospital characteristics, such as resources, organisational capacity, etc. for comparison studies;
δ) the assumption of multivariate normal distribution for random disturbances and the random parameters of relative effectiveness is not respected and no prior information on parameter distribution is available (Langford & Lewis, 1998; Marshall & Spiegelhalter, 2001).

The newly developed methodology proposed in this paper is an attempt to overcome the present restrictions for the general application of effectiveness studies.

2. The new methodology

The methodology we are proposing combines already existing methods and models in the field of the multivariate analysis and consists of a few steps, which solve the problems illustrated above.

α) First of all, if we interpret outcomes as latent constructs underlying a set of observable indicators (Gertler, 1988), they can be defined in statistical terms as latent variables (Muthen & Speckart, 1985; Mc Leod, 2001; Goldstein & Leyland, 2001). In order to obtain the solutions, we avoid traditional structural equation models, which lead to indeterminacy of latent scores (Vittadini, 1989). Instead the regression component decomposition (RCD) method (Schoenemann & Steiger, 1976), which approximates latent variables by means of linear combinations of their quantitative indicators, can be utilised. In particular, in order to obtain proxies of the latent healthcare outcomes when indicators are entirely qualitative or both qualitative and quantitative, RCD can be combined with methods of multidimensional scaling (Vittadini, 2001), such as Alsos Princals algorithm (De Leeuw & Van Rijckervorsel 1980; Young, 1981), which quantifies observable indicators and obtains their quantitative linear transformations according to principal components criteria in an iterative procedure.

β) Secondly, for these kinds of studies, the multivariate regression model is not adequate, as implicated in point β). Vittadini (2001) has proposed the combination of the multilevel model with the seemingly unrelated regression equations (SURE) model (Srivastava & Giles, 1987) in a unique model, entitled the SURE multilevel model. There are two reasons for this. First, the SURE model consists of a system of simultaneous equations, equal in number to the response variables, where disturbance terms related to different individuals in the same equation are non-correlated, but disturbance terms in separate equations are correlated. Second, the SURE model allows the specification of different regressors for each outcome, e.g. in the case of two different outcomes (final mental state of health and final physical state of health), we can indicate the initial state of health (which is different for each outcome) from among the various explicative variables.
In matrix notation, the SURE multilevel model specified for p outcomes can be expressed as:

$$\boldsymbol{y} = \boldsymbol{u}\boldsymbol{Z} + \boldsymbol{\beta}\boldsymbol{X} + \boldsymbol{e}, \tag{2}$$

where $\boldsymbol{y} = (\boldsymbol{y}'_1, \ldots, \boldsymbol{y}'_j, \ldots, \boldsymbol{y}'_p)$ is a row vector $(1 \times np)$, (with $\boldsymbol{y}'_j$, $j = 1, \ldots, p$, row vectors $(1 \times n)$); $\boldsymbol{u} = (\boldsymbol{u}'_1, \ldots, \boldsymbol{u}'_j, \ldots, \boldsymbol{u}'_p)$ is a row vector $(1 \times qp)$, (with $\boldsymbol{u}'_j$, $j = 1, \ldots, p$, row vectors $(1 \times q)$); $\boldsymbol{Z}$ is a 'design matrix' $(qp \times np)$; $\boldsymbol{\beta} = (\boldsymbol{\beta}'_1, \ldots, \boldsymbol{\beta}'_j, \ldots, \boldsymbol{\beta}'_p)$ is a row vector $(1 \times t)$ (with $\boldsymbol{\beta}'_j$, $j = 1, \ldots, p$, row vectors $(1 \times t_j)$, $t = \sum_{j=1}^{p} t_j$); $\boldsymbol{X}$ is a non-stochastic block diagonal matrix $(t \times np)$, (with typical element $\boldsymbol{X}_j$, $j = 1, \ldots, p$, non-null matrix $(t_j \times n)$); $\boldsymbol{e} = (\boldsymbol{e}'_1, \ldots, \boldsymbol{e}'_j, \ldots, \boldsymbol{e}'_p)$ is a row vector $(1 \times np)$, (with $\boldsymbol{e}'_j$, $j = 1, \ldots, p$, row vectors $(1 \times n)$).
Under the hypotheses of the SURE model, the stochastic vector of disturbances $\boldsymbol{e}$ follows a multivariate normal distribution with

$$\begin{aligned} E[\boldsymbol{e}] &= \boldsymbol{0}, \\ E[\boldsymbol{e}'\boldsymbol{e}] &= \boldsymbol{\Psi}_e = \boldsymbol{\Sigma}_e \otimes \boldsymbol{I}_n, \end{aligned} \tag{3}$$

where $\otimes$ denotes the Kronecker product (so that $\boldsymbol{\Psi}_e$ is $(np \times np)$) and

$$\boldsymbol{\Sigma}_e = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \ldots & \sigma_{1p} \\ \sigma_{21} & \sigma_2^2 & \ldots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \ldots & \sigma_p^2 \end{bmatrix}$$

is a $(p \times p)$ positive definite symmetric matrix. "The definition of $\boldsymbol{\Sigma}_e$ precludes the possibility of any linear dependencies among the contemporaneous disturbances in the p equations of the model" (Srivastava & Giles, 1987, p. 5).

γ) Moreover, with reference to point γ), we introduce another equation describing hospital characteristics. That means specifying a second level equation (i.e. referring to hospitals), which expresses the random parameters u_{jv} $(j = 1, \ldots, p; v = 1, \ldots, q)$ as a function of hospital characteristics

$$\boldsymbol{u} = \boldsymbol{\delta F} + \boldsymbol{m}, \tag{4}$$

where $\boldsymbol{\delta} = (\boldsymbol{\delta}'_1, \ldots, \boldsymbol{\delta}'_j, \ldots, \boldsymbol{\delta}'_p)$ is a row vector $(1 \times hp)$ containing fixed parameters for hospital covariates (with $\boldsymbol{\delta}'_j$, $j = 1, \ldots, p$, row vectors $(1 \times h)$); $\boldsymbol{F}$ is a non-stochastic block diagonal matrix $(hp \times qp)$ containing the observable characteristics of hospitals (with typical element $\boldsymbol{F}_j$, $j = 1, \ldots, p$, non-null matrix $(h \times q)$); $\boldsymbol{m} = (\boldsymbol{m}'_1, \ldots, \boldsymbol{m}'_j, \ldots, \boldsymbol{m}'_p)$ is a row vector $(1 \times qp)$ containing random residuals associated with hospitals (with $\boldsymbol{m}'_j$, $j = 1, \ldots, p$, row vectors $(1 \times q)$ and block diagonal covariance matrix $\boldsymbol{\Sigma}_m$—with typical element $s_j^2 \boldsymbol{I}_q$—as implied by the assumption of non-correlation between disturbances in separate equations and also between disturbances related to different individuals in the same equation). These random residuals indicate the relative effectiveness of hospitals (with respect to outcomes) adjusted for the characteristics of both patients and hospitals.
It is further assumed that

$$\boldsymbol{X}\boldsymbol{\Psi}_e^{-1}\boldsymbol{Z}' = \boldsymbol{O}, \tag{5}$$

where $\boldsymbol{O}$ is a matrix $(t \times np)$ of zeroes. This holds if values of explicative variables referring to the same hospital are standardised.

δ) The SURE multilevel model, like the multilevel model, is usually based on the assumption of multivariate normal distribution for both random disturbances and the random parameters of relative effectiveness (Srivastava & Giles, 1987; Montgomery, 1997). However, since this assumption is often too restrictive, in this paper we propose a generalisation of the SURE multilevel model based on the multivariate normal distribution of order λ, also known as 'exponential power distribution' (Subbotin, 1923; Box & Tiao, 1992). This distribution is symmetrical and flexible with respect to data and follows this formula:

$$f(\boldsymbol{y}) = k|\boldsymbol{\Sigma}_e^{-1}|^{1/2} \exp\left\{(-1)[(\boldsymbol{y} - \boldsymbol{uZ} - \boldsymbol{\beta X})\boldsymbol{\Psi}_e^{-1}(\boldsymbol{y} - \boldsymbol{uZ} - \boldsymbol{\beta X})']^{\lambda/2}\right\}, \tag{6}$$

where k is a coefficient of proportionality and $\lambda > 1$.
The model now satisfies the requirement of point δ).

With the inclusion of this last modification, we now have the complete structure of the general SURE multilevel model, the first statistical instrument capable of satisfying effectiveness study requirements in all their complexity.

3. A Bayesian approach to estimation

In order to estimate the model introduced, we propose a Bayesian approach, as did authors cited in Section 1, who have provided ranking among hospitals based on the multilevel model. "Recent computational advances, specifically in MCMC (Markov Chain Monte Carlo) methods, allow one to quantify the uncertainty associated with an institution's rank and so to determine the extent to which conclusions may be based on explicit rankings" (Marshall & Spiegelhalter, 2001, p. 128).

Under the general assumption of uniform non-informative distribution for 'random effects' u_{jv} ($j = 1, \ldots, p; v = 1, \ldots, q$) (valid when we do not have sufficient information to hypothesise an informative prior distribution), the model described by (2) and (4) is interpretable as a mixed model, where the parameters $\beta_{j\tau}$ ($j = 1, \ldots, p; \tau = 1, \ldots, t_j$) are 'fixed effects' and are *exchangeable* with each other and with respect to u_{jv} (Gelman *et al*., 1995, pp. 368–369; Lindley & Smith, 1972; Smith, 1973). Before going ahead, however, it is necessary to clarify the two new terms introduced.

1. "The terms 'fixed' and 'random' come from the non-Bayesian statistical tradition and are somewhat confusing in a Bayesian context, where all unknown parameters are treated as 'random'. The non-Bayesian view considers 'fixed effects' to be fixed unknown quantities, but the standard procedures proposed to estimate these parameters based on specific repeated-sampling properties, happen to be equivalent to the Bayesian posterior inference under a non-informative (uniform) prior distribution" (Gelman *et al*., 1995, pp. 368–369). Therefore, the fixed effects are interpretable in a Bayesian context as random parameters of a conventional non-informative uniform prior distribution.
2. The definition of *exchangeability* is given by Gelman *et al*. (1995) on pp. 123–124. "If no information—other than the data $\boldsymbol{y}$—is available to distinguish any of the parameter vectors $\boldsymbol{u}_j$'s from any of the others, and no ordering or grouping of the parameters can be made, one must assume symmetry among the parameters in their prior distribution. This symmetry is represented probabilistically by *exchangeability*; the parameters $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_j, \ldots, \boldsymbol{u}_p)$ are *exchangeable* in their joint distribution if $P(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_j, \ldots, \boldsymbol{u}_p)$ is invariant to permutations of the indexes $(1, \ldots, p)$. The simplest form of an *exchangeable* distribution has each of the parameter vectors $\boldsymbol{u}_j$ as an independent sample from a population distribution governed by some unknown parameter vector $\boldsymbol{\phi}$":

$$P(\boldsymbol{u}|\boldsymbol{\phi}) = \prod_{j=1}^{p} P(\boldsymbol{u}_j|\boldsymbol{\phi}). \tag{7}$$

Coming back to the model introduced, the likelihood function is formulated from (6) as:

$$P(\boldsymbol{y}|\boldsymbol{u}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_e^{-1}) = k|\boldsymbol{\Sigma}_e^{-1}|^{n/2} \exp\left\{(-1)[(\boldsymbol{y} - \boldsymbol{u}\boldsymbol{Z} - \boldsymbol{\beta}\boldsymbol{X})\boldsymbol{\Psi}_e^{-1}(\boldsymbol{y} - \boldsymbol{u}\boldsymbol{Z} - \boldsymbol{\beta}\boldsymbol{X})']^{\lambda/2}\right\}, \tag{8}$$

and assuming independence between the prior information regarding the elements of $\boldsymbol{u}$, $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}_e^{-1}$, and following Jeffreys' invariance theory (Zellner, 1971, p. 242; Srivastava & Giles, 1987, p. 317), the joint non-informative prior distribution for $(\boldsymbol{u}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_e^{-1})$ is given by:

$$P(\boldsymbol{u}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_e^{-1}) = P(\boldsymbol{u})P(\boldsymbol{\beta})P(\boldsymbol{\Sigma}_e^{-1}), \tag{9}$$

where:

1) $P(\boldsymbol{u}) \propto$ uniform $(-\infty < \boldsymbol{u} < +\infty)$; $P(u_{jv}) \propto$ uniform $(-\infty < u_{jv} < +\infty)$,
2) $P(\boldsymbol{\beta}) \propto$ uniform $(-\infty < \boldsymbol{\beta} < +\infty)$; $P(\beta_{j\tau}) \propto$ uniform $(-\infty < \beta_{j\tau} < +\infty)$,
3) $P(\boldsymbol{\Sigma}_e^{-1}) \propto |\boldsymbol{\Sigma}_e^{-1}|^{-(p+1)/2}$ (prior on the $(p+1)/2$ distinct elements of $\boldsymbol{\Sigma}_e^{-1}$).

On combining (8) and (9), the joint posterior distribution for $(\boldsymbol{u}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_e^{-1})$ is

$$\begin{aligned} P(\boldsymbol{u}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_e^{-1}|\boldsymbol{y}) &\propto P(\boldsymbol{y}|\boldsymbol{u}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_e^{-1})P(\boldsymbol{u}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_e^{-1}) \\ &\propto k|\boldsymbol{\Sigma}_e^{-1}|^{[n-(p+1)]/2} \exp\left\{(-1)[(\boldsymbol{y} - \boldsymbol{u}\boldsymbol{Z} - \boldsymbol{\beta}\boldsymbol{X})\boldsymbol{\Psi}_e^{-1}(\boldsymbol{y} - \boldsymbol{u}\boldsymbol{Z} - \boldsymbol{\beta}\boldsymbol{X})']^{\lambda/2}\right\}. \end{aligned} \tag{10}$$

For SURE models with multivariate normal distribution for the disturbance vector $\boldsymbol{e}$ and non-informative prior distribution for $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}_e^{-1}$, the conditional posterior distribution for $\boldsymbol{\beta}$ given $\boldsymbol{\Sigma}_e^{-1}$

is multivariate normal in form (Zellner, 1971, p. 242; Srivastava & Giles, 1987, p. 317). In Appendix A, we will demonstrate that for the SURE multilevel model, introduced in Section 2, the conditional posterior density for $\boldsymbol{u}$, given $\boldsymbol{\Sigma}_e^{-1}$ (indicated by $P(\boldsymbol{u}|\boldsymbol{\Sigma}_e^{-1}, \boldsymbol{y})$) and the conditional posterior density for $\boldsymbol{\beta}$, given $\boldsymbol{\Sigma}_e^{-1}$ (indicated by $P(\boldsymbol{\beta}|\boldsymbol{\Sigma}_e^{-1}, \boldsymbol{y})$) are in the multivariate normal form of order λ.

If there is no information about $\boldsymbol{\Sigma}_e^{-1}$, Zellner (1971) suggests replacing $\boldsymbol{\Sigma}_e^{-1}$ by a consistent estimate $\boldsymbol{S}_e^{-1}$, which yields the conditional posterior sample densities $P(\boldsymbol{u}|\boldsymbol{S}_e^{-1}, \boldsymbol{y})$ and $P(\boldsymbol{\beta}|\boldsymbol{S}_e^{-1}, \boldsymbol{y})$. "In large samples $\boldsymbol{S}_e^{-1}$ will not differ markedly from $\boldsymbol{\Sigma}_e^{-1}$, and thus going ahead with the assumption $\boldsymbol{\Sigma}_e^{-1} = \boldsymbol{S}_e^{-1}$ will produce satisfactory results" (Zellner, 1971, p. 243). In small samples, however, it is better to obtain the marginal posterior densities $P(\boldsymbol{u}|\boldsymbol{y})$ and $P(\boldsymbol{\beta}|\boldsymbol{y})$. "In finite samples the substitution of $\boldsymbol{S}_e^{-1}$ for $\boldsymbol{\Sigma}_e^{-1}$ may be less appealing. In fact, one of the principal advantages of the Bayesian approach to the estimation is that no such substitution is necessary. Instead of basing our inferences on the conditional posterior density, the elements of $\boldsymbol{\Sigma}_e^{-1}$ may be treated as 'nuisance parameters', and eliminated from by multivariate integration to yield the marginal posterior density" (Srivastava & Giles, 1987, p. 318).

Once the posterior distributions for $\boldsymbol{u}$ and $\boldsymbol{\beta}$ are obtained, expected values and credible intervals for fixed parameters $\boldsymbol{\beta}$ provide information regarding the relationship between patient characteristics and outcomes. Next, with reference to (4), regressing the expected values for random parameters $\boldsymbol{u}$ on matrix $\boldsymbol{F}$, we obtain estimates for fixed parameters $\boldsymbol{\delta}_j$ (which provide information regarding the relationship between hospital characteristics and outcomes) and expected values and credible intervals for random residuals $\boldsymbol{m}_j$. An effectiveness ranking of hospitals is provided by comparing the interval boundaries of $\boldsymbol{m}_j$. We can say that hospital v_1 can be considered as better than hospital v_2 when the lower bound of credible interval in reference to v_1 is greater than the upper bound of v_2. In the opposite case, hospitals v_1 and v_2 demonstrate the same level of effectiveness (Goldstein & Spiegelhalter, 1996). If, as usually happens in evaluation studies, the measurement of relative effectiveness is repeated at time intervals with regularity, information drawn at a previous time can be utilised in order to improve the estimates at the following interval. The posterior distribution at the previous time becomes the prior distribution at the following interval.

4. The application

We have applied the proposed methodology in an effectiveness study for a set of public hospitals in four Lombardy health districts (Milan, Como, Lecco, Brescia). Two diagnosis related groups (DRGs) were considered regarding two chronic-degenerative clinical conditions: Cardiac Arrhythmia and Conduction Disorders with CC (DRG 138), and Medical Back Problems (DRG 243). During the period July 1998–July 1999, the SF-12 Quality of Life questionnaire (a 12-item questionnaire which produces separate mental and physical health status scores) was administered to a sample of patients previously discharged from hospital and coded by DRG 138 or 243.

A second survey was carried out 1 year later (October 1999–November 2000) on re-hospitalised patients who had previously answered the first questionnaire, adopting the same SF-12 Quality of Life questionnaire. The re-screened sample resulted in a total of 1213 interviews.

By means of the method described in Section 2 point α), from the twelve items in the SF-12 questionnaire, we obtain two latent outcomes: mental quality of life (state of mental health) and physical quality of life (state of physical health). In reference to these outcomes, the aim of the analysis is the evaluation of relative effectiveness of participating hospitals on the re-hospitalised patients, adjusted for patient and hospital characteristics.

Variables utilised in the study are described as follows.

Dependent variables, i.e. outcomes, are classified as: y_1—outcome of patient's final state of mental health (continuous); y_2—outcome of patient's final state of physical health (continuous).

Patient-specific covariates are classified as: x_1—outcome of patient's initial state of mental health (continuous); x_2—outcome of patient's initial state of physical health (continuous); x_3—gender (dichotomous); x_4—length of stay (number of days; dichotomous); x_5, x_6—first and second re-hospitalisation (dichotomous), x_7—age (number of years; discrete).

Hospital-specific covariates are classified as: f_1, f_2, f_3—hospital type: case-mix (relative level of case complexity in a hospital with respect to the regional mean; continuous); number of operating rooms (discrete); average cost of hospital stay (continuous); f_4—efficiency: revenue/costs ratio (continuous); f_5—structural dimensions: number of beds (discrete); f_6, f_7—personnel dimensions: number of physicians (discrete); number of other personnel (discrete).

From the computational point of view, following Goldstein & Spiegelhalter (1996) and Spiegelhalter *et al*. (1996), an SAS procedure based on MCMC algorithm (Gibbs sampling with rejection/acceptance steps) has been developed in order to obtain numerical solutions for the conditional posterior distributions $P(\boldsymbol{\beta}|\boldsymbol{S}_e^{-1}, \boldsymbol{y})$ and $P(\boldsymbol{u}|\boldsymbol{S}_e^{-1}, \boldsymbol{y})$, which are in the multivariate normal form of order $\lambda = 4$. (Note that the substitution of $\boldsymbol{S}_e^{-1}$ for $\boldsymbol{\Sigma}_e^{-1}$ is adequate, since the sample size corresponds to the 10% of the population size). As an example, Fig. 1 shows that, for physical outcomes, the theoretical cumulate density function for random parameter u_{jv}, regarding the hospital codified as 30156, is nearly superimposed on the empirical cumulate density function; this holds for each hospital and each outcome.

With reference to DRG 138, Table B1 shows the expected values, the standard deviations and the credible intervals for parameters $\beta_{j\tau}$ (regarding patient characteristics) and δ_{jw} (regarding hospital characteristics); Fig. B2 represents the credible intervals for random residuals m_{jv} of mental and physical outcomes. Finally, Table B2 shows the results of hospital comparisons (obtained by means of credible intervals for m_{jv}) and the ranking of hospitals.

Tables B3 and B4 and Fig. B3 show analogous results for DRG 243. (See Tables B1–B4 and Figs B2–B3 in Appendix B.)

In Tables B1 and B3, we observe that the coefficients of patient-specific covariates are negligible, which indicates that we have patients who are similar in age and gravity of initial state of health (chronic

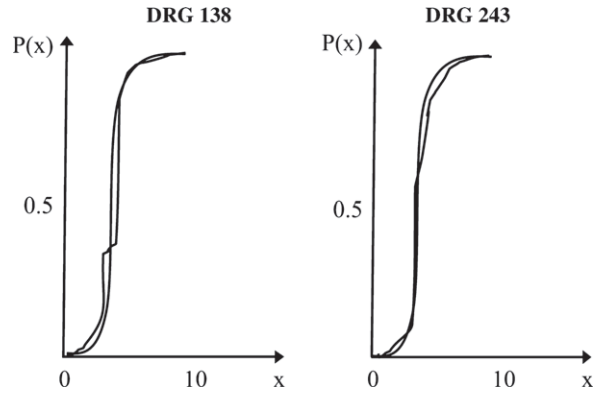

FIG. 1. Theoretical and empirical cumulate density function for random parameter u_{jv} regarding the hospital codified as 30156—physical outcomes DRG 138 and DRG 243.

patients). On the contrary, coefficients of hospital-specific covariates are more relevant. In particular, we note that coefficients of the same covariates differ between the two considered DRGs and also between the mental and physical outcomes for the same DRG. It is therefore clear that hospitals obtaining better results regarding mental outcomes have different characteristics than hospitals which obtain better results for the physical outcomes.

In particular, for DRG 138, the hospitals with the best results for mental and physical outcomes seem to utilise more nursing personnel and less doctors. These findings can be explained by the fact that chronic and long-term diseases need more regular care than intensive and specialised interventions. Moreover, mental outcomes are negatively related to the number of beds and positively to the case-mix, while physical outcomes depend positively on the number of beds and negatively on the case-mix. In fact, physical aspects involved in DRG 138 need no sub-specialised treatments (i.e. cardiac intensive care unit), which can be treated also in general hospitals; instead, mental aspects involved in DRG 138 are connected with particular treatments, which can be treated only in small specialised hospitals.

For the physical aspects involved in DRG 243, the best results are obtained by hospitals offering long-term care with more nursing personnel. In fact, the best hospitals are very strongly and positively associated with a higher number of nursing personnel and doctors and negatively with the number of beds. The negativity of the coefficients regarding case-mix also suggests phenomena of case-mix selection. Differently, the mental aspects of DRG 243 outcomes are positively dependent on the case-mix and the number of nursing personnel, but negatively on the number of doctors and number of beds.

Regarding hospital comparisons (based on the criteria indicated at the end of the previous section), groups of hospitals characterised by different grades of effectiveness are clearly identified (see Tables B2 and B4 and Figs B2 and B3). In synthesis, there are three groups of hospitals characterised as follows: ‘above the average’, ‘on the average’ and ‘below the average’ (note that the average level of effectiveness equals zero), and each hospital is associated with a different level of variability (see Figs B2 and B3). In particular, a large credible interval means that there is a different grade of effectiveness for different groups of patients in the same hospital, i.e. the hospital in question has different healthcare capacities regarding different states of health.

The results of the comparisons differ between the mental and physical outcomes.

In particular, for the mental aspects involved in DRG 138, most part of the hospitals considered in the analysis indicate levels of effectiveness which are ‘on the average’ or ‘above the average’; in particular, hospital 30914 indicates the best level of effectiveness and therefore can be qualified as ‘excellent’. Instead, some attention must be devoted to the hospital 30024, characterised by a level of effectiveness which is considerably ‘below the average’. Moreover, with reference to the hospitals characterised by large credible intervals, a case by case study is required, as illustrated previously. For the physical outcomes, a different ranking of hospitals is identified; in particular, most part of hospitals involved in the analysis indicate a lower grade of effectiveness with respect to the mental outcomes, except for hospitals 30156 and 30024. Some attention must be devoted to hospital 30274 and also to hospitals 30024 and 30902.

For the mental aspects involved in DRG 243, the hospitals characterised by the largest levels of effectiveness (namely the last five of the sequence) indicate also the largest credible intervals, requiring therefore a case by case study. The remaining hospitals, except for hospital 30022, are characterised by small credible intervals and indicate different levels of effectiveness. Some attention must be devoted to hospitals 30022, 30157 and 30159. For the physical outcomes, again the last five hospitals of the sequence are characterised by the largest levels of effectiveness, but indicate different levels of effectiveness with respect to the mental outcomes. Note that in this case, hospitals 30157 and 30159 indicate levels of effectiveness which are ‘above the average’.

5. Conclusions and further research

In this paper we have proposed a new methodology for the measurement of relative effectiveness of hospital care under general conditions. For this topic, simultaneous dependency of outcomes on covariates, hospital-specific covariates, multivariate normal distribution of order λ and non-informative prior distribution have been assumed.

Future developments regarding three possible extensions of the methodology proposed are:

1. In order to generalise the original SURE multilevel model, we have introduced the assumption of multivariate normal distribution of order λ for random disturbances and the random parameters of relative effectiveness. A further generalisation would be the assumption of a non-symmetrical distribution, e.g. the distribution entitled 'Stable Alpha' (Samorodmitsky & Takku, 1994).
2. Alternatively, we suggest the introduction of more general models, like the Cluster Weighted Models (Gershenfeld *et al*., 1999), which are characterised by higher predictive capacity, without sacrificing interpretability.
3. The methodology proposed has been applied to observational data. In order to attribute a general value to the results presented in this paper, a possible extension would be to implement this methodology using experimental data, drawn from experimental or quasi-experimental designs typical of epidemiological and clinical studies. However, while epidemiological and clinical outcomes and related case-mix variables can measure the effectiveness of the health institutions more precisely (Kerr *et al.*, 1998), problems of cost and time and the lack of adequate personnel can make their use very problematic.

Acknowledgements

The authors would like to express their thanks to the referees for their valuable remarks and comments.

References

ALBERT, J. & CHIB, S. (1997) Bayesian tests and model diagnostics in conditionally independent hierarchical models. *J. Am. Stat. Assoc.*, **92**, 916–925.

BASILEVSKY, A. (1983) *Applied Matrix Algebra in the Statistical Sciences*. Amsterdam: North Holland.

BOX, G. E. P. & TIAO, G. C. (1992) *Bayesian Inference in Statistical Analysis*. Chichester: Wiley.

DE LEEUW, J. & VAN RIJCKERVORSEL, J. (1980) Homals and Princals, some generalizations of components analysis. *Data Analysis and Informatics*, **2**, 231–241.

GELMAN, A., CARLIN, J. B., STERN, H. S. & RUBIN, D. B. (1995) *Bayesian Data Analysis*. London: Chapman & Hall.

GERSHENFELD, N. A., SCHONER, B. & METOIS, E. (1999) Cluster-weighted modelling for time-series analysis. *Nature*, **397**, 329–332.

GERTLER, P. J. (1988) A latent-variable model of quality determination. *J. Bus. Econ. Stat.*, **9**, 241–252.

GOLDSTEIN, H. & LEYLAND, A. H. (2001) Further developments. *multilevel modelling of Health Statistics* (H. Goldstein & A. H. Leyland eds). Chichester: Wiley, pp. 175–186.

GOLDSTEIN, H. & SPIEGELHALTER, D. J. (1996) League tables and their limitations: statistical issues in comparison of institutional performance. *J. R. Stat. Soc. Ser. A*, **159**, 385–443.

KERR, E. A., MC GLYNN, E. A. & DAMBERG, C. L. (1998) Developing and using a clinical information system. *Health Information Systems: Design Issues and Analytic Application* (E. A. Mc Glynn, C. L. Damberg, E. A. Kerr & R. H. Brook eds). RAND Health Corporation, http://www.rand.org/publications/MR/MR967/MR967.pdf, pp. 141–187.

LANGFORD, I. H. & LEWIS, T. (1998) Outliers in multilevel data. *J. R. Stat. Soc. Ser. A*, **161**, 121–160.

LEYLAND, A. H. & BODDY, A. (1998) League tables and acute myocardial infarction. *Lancet*, **351**, 555–558.

LINDLEY, D. V. & SMITH, A. F. M. (1972) Bayes estimates for the linear model (with discussion). *J. R. Stat. Soc. Ser. B Meth.*, **34**, 1–41.

MARSHALL, E. C. & SPIEGELHALTER, D. J. (2001) Institutional performance. *multilevel modelling of Health Statistics* (H. Goldstein & A. H. Leyland eds). Chichester: Wiley, pp. 127–142.

MC LEOD, A. (2001) Multivariate multilevel models. *Multilevel Modelling of Health Statistics* (H. Goldstein & A. H. Leyland eds). Chichester: Wiley, pp. 59–74.

MONTGOMERY, A. L. (1997) Hierarchical Bayes models for micro-marketing strategies. *Case Studies in Bayesian Statistics*, Vol. 3 (C. Gatsonis *et al*. eds). New York: Springer, pp. 95–141.

MORRIS, C. N. & CHRISTIANSEN, C. L. (1996) Hierarchical models for ranking and identifying extremes, with applications. *Bayesian Statistics 5* (J. O. Bernardo, J. O. Berger, A. P. Dawid & A. F. M. Smith eds). Oxford: Oxford University Press, 277–297.

MUTHEN, B. & SPECKART, G. (1985) Latent variable probit ANCOVA: treatment effects in the California Civil Addict Programme. *Br. J. Math. Stat. Psychol.*, **38**, 161–170.

NORMAND, S. L. T., GLICKMAN, M. E. & RYAN, T. J. (1995) Modelling mortality rates for elderly heart attack patients: profiling hospitals in the co-operative cardiovascular project. *Case Studies in Bayesian Statistics*, Vol. 2 (C. Gatsonis *et al.* eds). New York: Springer, pp. 435–456.

OPIT, L. J. (1993) The measurement of health service outcomes. *Oxford Textbook of Health Care*, vol. 10. London: OLJ.

RICE, N. & LEYLAND, A. (1996) Multilevel models: application to health data. *J. Health Serv. Res. Policy*, **1**, 154–164.

SAMORODMITSKY, G. & TAKKU, M. S. (1994) *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. London: Chapman & Hall.

SCHOENEMANN, P. H. & STEIGER, J. (1976) Regression component analysis. *Br. J. Math. Stat. Psychol.*, **29**, 175–189.

SMITH, A. F. M. (1973) A general Bayesian linear model. *J. R. Stat. Soc. Ser. B Meth.*, **35**, 67–75.

SPIEGELHALTER, D., THOMAS, A., BEST, N. & GILKS, W. (1996) *Bayesian Using Gibbs Sampling Manual*. Cambridge: MRC Biostatistic Unit, Institute of Public Health.

SRIVASTAVA, V. K. & GILES, D. E. A. (1987) *Seemingly Unrelated Regression Equations Models*. New York: Marcel Dekker.

SUBBOTIN, M. T. (1923) On the law of frequency of errors. *Matem. Sbornik*, **31**, 296–301.

THOMAS, N., LONGFORD, N. T. & ROLPH, J. E. (1994) Empirical Bayes methods for estimating hospital-specific mortality rates. *Stat. Med.*, **13**, 889–903.

VITTADINI, G. (1989) Indeterminacy problems in the Lisrel model. *Multivariate Behav. Res.*, **24**, 397–414.

VITTADINI, G. (2001) On the use of multivariate regression models in the context of multilevel analysis. *Advances in Classification and Data Analysis* (S. Borra, M. Rocci & M. Schader eds). Heidelberg: Springer, pp. 225–232.

YOUNG, F. (1981) Quantitative analysis of qualitative data. *Psychometrika*, **46**, 357–388.

ZASLAVSKY, A. (2001) Statistical issues in reporting quality data: small samples and casemix variation. *Int. J. Qual. Health Care*, **13**, 481–488.

ZELLNER, A. (1971) *An Introduction to Bayesian Inference in Econometrics*. Chicester: Wiley.

Appendix A. Multivariate normality of order λ for $P(\boldsymbol{u}|\boldsymbol{\Sigma}_e^{-1}, \boldsymbol{y})$ and $P(\boldsymbol{\beta}|\boldsymbol{\Sigma}_e^{-1}, \boldsymbol{y})$

We demonstrate that for the SURE multilevel model, introduced in Section 2, the conditional posterior distributions for $\boldsymbol{u}$ and $\boldsymbol{\beta}$, (indicated by $P(\boldsymbol{u}|\boldsymbol{\Sigma}_e^{-1}, \boldsymbol{y})$ and $P(\boldsymbol{\beta}|\boldsymbol{\Sigma}_e^{-1}, \boldsymbol{y})$, respectively), are in the multivariate normal form of order λ.

For models characterised by correlated disturbance terms in separate equations (like the SURE model), as initial estimates for $\boldsymbol{u}$ and $\boldsymbol{\beta}$ (indicated by $\hat{\boldsymbol{u}}$ and $\hat{\boldsymbol{\beta}}$, respectively), the Generalised Least Squares estimators are utilised, which in this case correspond to:

$$[\hat{\boldsymbol{u}}, \hat{\boldsymbol{\beta}}] = [\boldsymbol{y}\boldsymbol{\Psi}_e^{-1}\boldsymbol{Z}'(\boldsymbol{Z}\boldsymbol{\Psi}_e^{-1}\boldsymbol{Z}')^{-1}, \boldsymbol{y}\boldsymbol{\Psi}_e^{-1}\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{\Psi}_e^{-1}\boldsymbol{X}')^{-1}], \tag{A.1}$$

given that assumption (5) implies:

$$\boldsymbol{V} = \begin{pmatrix} \boldsymbol{Z}\boldsymbol{\Psi}_e^{-1}\boldsymbol{Z}' & \boldsymbol{Z}\boldsymbol{\Psi}_e^{-1}\boldsymbol{X}' \\ \boldsymbol{X}\boldsymbol{\Psi}_e^{-1}\boldsymbol{Z}' & \boldsymbol{X}\boldsymbol{\Psi}_e^{-1}\boldsymbol{X}' \end{pmatrix} = \begin{pmatrix} \boldsymbol{Z}\boldsymbol{\Psi}_e^{-1}\boldsymbol{Z}' & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{X}\boldsymbol{\Psi}_e^{-1}\boldsymbol{X}' \end{pmatrix} = \begin{pmatrix} \boldsymbol{V}_{\boldsymbol{u}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{V}_{\boldsymbol{\beta}} \end{pmatrix}. \tag{A.2}$$

The joint posterior distribution in (10) can be rewritten as a function of initial estimates $\hat{\boldsymbol{u}}$ and $\hat{\boldsymbol{\beta}}$ as:

$$\begin{aligned} P(\boldsymbol{u}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_e^{-1}|\boldsymbol{y}) \propto k|\boldsymbol{\Sigma}_e^{-1}|^{[n-(p+1)]/2} \\ \times \exp\{(-1)\{[(\boldsymbol{y} - \hat{\boldsymbol{u}}\boldsymbol{Z} - \hat{\boldsymbol{\beta}}\boldsymbol{X}) - (\boldsymbol{u} - \hat{\boldsymbol{u}})\boldsymbol{Z} - (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})\boldsymbol{X}] \\ \times \boldsymbol{\Psi}_e^{-1}[(\boldsymbol{y} - \hat{\boldsymbol{u}}\boldsymbol{Z} - \hat{\boldsymbol{\beta}}\boldsymbol{X}) - (\boldsymbol{u} - \hat{\boldsymbol{u}})\boldsymbol{Z} - (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})\boldsymbol{X}]'\}^{\lambda/2}\}. \end{aligned} \tag{A.3}$$

On indicating by $\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3$ the quantities:

$$\begin{aligned} \boldsymbol{q}_1 &= \boldsymbol{y} - \hat{\boldsymbol{u}}\boldsymbol{Z} - \hat{\boldsymbol{\beta}}\boldsymbol{X}, \\ \boldsymbol{q}_2 &= (\boldsymbol{u} - \hat{\boldsymbol{u}})\boldsymbol{Z}, \\ \boldsymbol{q}_3 &= (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})\mathbf{X}, \end{aligned} \tag{A.4}$$

and noting that (5), (A.1) and (A.2) imply:

$$\begin{aligned} \boldsymbol{q}_1\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_2' &= 0, \\ \boldsymbol{q}_1\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_3' &= 0, \\ \boldsymbol{q}_2\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_3' &= 0, \end{aligned} \tag{A.5}$$

we can rewrite formula (A.3) as:

$$P(\boldsymbol{u}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_e^{-1}|\boldsymbol{y}) \propto k|\boldsymbol{\Sigma}_e^{-1}|^{[n-(p+1)]/2} \exp\{(-1)[\boldsymbol{q}_1\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_1' + \boldsymbol{q}_2\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_2' + \boldsymbol{q}_3\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_3']^{\lambda/2}\}. \tag{A.6}$$

Given that a quadratic form (which is equal to a scalar) is equivalent to its trace, the (A.6) is equivalent to:

$$\left\{\mathrm{tr}[\boldsymbol{q}_1\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_1' + \boldsymbol{q}_2\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_2' + \boldsymbol{q}_3\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_3']\right\}^{\lambda/2}. \tag{A.7}$$

Applying to formula (A.7) the property of the trace of the matrix sum, namely $\mathrm{tr}(\boldsymbol{A}+\boldsymbol{B}) = \mathrm{tr}(\boldsymbol{A}) + \mathrm{tr}(\boldsymbol{B})$ (Basilevsky, 1983, p. 103), we have:

$$\left\{\mathrm{tr}[\boldsymbol{q}_1\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_1'] + \mathrm{tr}[\boldsymbol{q}_2\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_2'] + \mathrm{tr}[\boldsymbol{q}_3\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_3']\right\}^{\lambda/2}, \tag{A.8}$$

which can be factorised as:

$$\left\{\text{tr}[\boldsymbol{q}_1\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_1'] + \text{tr}[\boldsymbol{q}_2\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_2'] + \text{tr}[\boldsymbol{q}_3\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_3']\right\}^2 \times \left\{\text{tr}[\boldsymbol{q}_1\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_1'] + \text{tr}[\boldsymbol{q}_2\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_2'] + \text{tr}[\boldsymbol{q}_3\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_3']\right\}^{\lambda/2-2}. \tag{A.9}$$

On completing the square in (A.9), we have:

$$\begin{aligned}&\text{tr}[\boldsymbol{q}_1\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_1'] \cdot \text{tr}[\boldsymbol{q}_1\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_1'] + \text{tr}[\boldsymbol{q}_2\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_2'] \cdot \text{tr}[\boldsymbol{q}_2\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_2'] \\ &\quad + \text{tr}[\boldsymbol{q}_3\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_3'] \cdot \text{tr}[\boldsymbol{q}_3\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_3'] + 2\,\text{tr}[\boldsymbol{q}_1\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_1'] \cdot \text{tr}[\boldsymbol{q}_2\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_2'] \\ &\quad + 2\,\text{tr}[\boldsymbol{q}_1\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_1'] \cdot \text{tr}[\boldsymbol{q}_3\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_3'] + 2\,\text{tr}[\boldsymbol{q}_2\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_2'] \cdot \text{tr}[\boldsymbol{q}_3\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_3'].\end{aligned} \tag{A.10}$$

The first three terms in (A.10) can be rewritten as:

$$[\boldsymbol{q}_1\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_1']^2 + [\boldsymbol{q}_2\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_2']^2 + [\boldsymbol{q}_3\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_3']^2. \tag{A.11}$$

The second three terms in (A.10) can be rewritten as:

$$\begin{aligned}&2\,\text{tr}[\boldsymbol{q}_1\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_1'] \cdot \text{tr}[\boldsymbol{q}_2\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_2'\boldsymbol{\Psi}_e\boldsymbol{\Psi}_e^{-1}] + 2\,\text{tr}[\boldsymbol{q}_1\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_1'] \cdot \text{tr}[\boldsymbol{q}_3\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_3'\boldsymbol{\Psi}_e\boldsymbol{\Psi}_e^{-1}] \\ &\quad + 2\,\text{tr}[\boldsymbol{q}_2\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_2'] \cdot \text{tr}[\boldsymbol{q}_3\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_3'\boldsymbol{\Psi}_e\boldsymbol{\Psi}_e^{-1}].\end{aligned} \tag{A.12}$$

Applying to the (A.12) the commutative law of the matrix product, namely $\text{tr}(\boldsymbol{AB}) = \text{tr}(\boldsymbol{BA})$ (Basilevsky, 1983, p. 103), we have:

$$\begin{aligned}&2\,\text{tr}[\boldsymbol{q}_1\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_1'] \cdot \text{tr}[\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_2\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_2'\boldsymbol{\Psi}_e] + 2\,\text{tr}[\boldsymbol{q}_1\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_1'] \cdot \text{tr}[\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_3\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_3'\boldsymbol{\Psi}_e] \\ &\quad + 2\,\text{tr}[\boldsymbol{q}_2\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_2'] \cdot \text{tr}[\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_3\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_3'\boldsymbol{\Psi}_e].\end{aligned} \tag{A.13}$$

Using the property of the trace of the Kronecker product, namely $\text{tr}(\boldsymbol{A} \otimes \boldsymbol{B}) = \text{tr}(\boldsymbol{A}) \cdot \text{tr}(\boldsymbol{B})$ (Basilevski, 1983, p. 104), we obtain:

$$\begin{aligned}&2\,\text{tr}[\boldsymbol{q}_1\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_1' \otimes \boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_2\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_2'\boldsymbol{\Psi}_e] + 2\,\text{tr}[\boldsymbol{q}_1\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_1' \otimes \boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_3\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_3'\boldsymbol{\Psi}_e] \\ &\quad + 2\,\text{tr}[\boldsymbol{q}_2\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_2' \otimes \boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_3\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_3'\boldsymbol{\Psi}_e].\end{aligned} \tag{A.14}$$

Note that the first term of each of the three Kronecker products is equal to a scalar; recall that for a scalar k, we have $k \otimes \boldsymbol{A} = k \cdot \boldsymbol{A}$ (Basilevski, 1983, p. 87), from the (A.14) we obtain:

$$\begin{aligned}&2\,\text{tr}[\boldsymbol{q}_1\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_1' \cdot \boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_2\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_2'\boldsymbol{\Psi}_e] + 2\,\text{tr}[\boldsymbol{q}_1\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_1' \cdot \boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_3\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_3'\boldsymbol{\Psi}_e] \\ &\quad + 2\,\text{tr}[\boldsymbol{q}_2\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_2' \cdot \boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_3\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_3'\boldsymbol{\Psi}_e] = 0.\end{aligned} \tag{A.15}$$

Applying an analogous procedure to the $(\lambda/2 - 2)$ powers in (A.9) and given that the 'mixed products' always equal to zero, it is easy to demonstrate that formula (A.6) is equivalent to:

$$P(\boldsymbol{u}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_e^{-1}|\boldsymbol{y}) \propto k|\boldsymbol{\Sigma}_e^{-1}|^{[n-(p+1)]/2} \times \exp\left\{(-1)\{[\boldsymbol{q}_1\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_1']^{\lambda/2} + [\boldsymbol{q}_2\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_2']^{\lambda/2} + [\boldsymbol{q}_3\boldsymbol{\Psi}_e^{-1}\boldsymbol{q}_3']^{\lambda/2}\}\right\}. \tag{A.16}$$

On substituting the definitions for $\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3$, the (A.16) can be rewritten as:

$$\begin{aligned} P(\boldsymbol{u}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_e^{-1}|\boldsymbol{y}) \propto\ & c|\boldsymbol{\Sigma}_e^{-1}|^{[n-(p+1)+q+t]/2} \exp\left\{(-1)[(\boldsymbol{y} - \hat{\boldsymbol{u}}\boldsymbol{Z} - \hat{\boldsymbol{\beta}}\boldsymbol{X})\boldsymbol{\Psi}_e^{-1}(\boldsymbol{y} - \hat{\boldsymbol{u}}\boldsymbol{Z} - \hat{\boldsymbol{\beta}}\boldsymbol{X})']^{\lambda/2}\right\} \\ & \times c|\boldsymbol{\Sigma}_e^{-1}|^{-q/2} \exp\left\{(-1)[(\boldsymbol{u} - \hat{\boldsymbol{u}})\boldsymbol{V}_{\boldsymbol{u}}(\boldsymbol{u} - \hat{\boldsymbol{u}})']^{\lambda/2}\right\} \\ & \times c|\boldsymbol{\Sigma}_e^{-1}|^{-t/2} \exp\left\{(-1)[(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})\boldsymbol{V}_{\boldsymbol{\beta}}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})']^{\lambda/2}\right\}, \end{aligned} \tag{A.17}$$

where $c = \sqrt[3]{k}$ is a proportionality coefficient.

The (A.17) expresses the joint posterior distribution for parameters $(\boldsymbol{u}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_e^{-1})$ as a product of three densities, which are proportional to $P(\boldsymbol{\Sigma}_e^{-1}|\boldsymbol{y})$, $P(\boldsymbol{u}|\boldsymbol{\beta}, \boldsymbol{\Sigma}_e^{-1}, \boldsymbol{y})$ and $P(\boldsymbol{\beta}|\boldsymbol{\Sigma}_e^{-1}, \boldsymbol{y})$, respectively. From (A.17) we can see that according to the *exchangeability* for $\boldsymbol{u}$ and $\boldsymbol{\beta}$, $P(\boldsymbol{u}|\boldsymbol{\beta}, \boldsymbol{\Sigma}_e^{-1}, \boldsymbol{y}) = P(\boldsymbol{u}|\boldsymbol{\Sigma}_e^{-1}, \boldsymbol{y})$, and distributions $P(\boldsymbol{u}|\boldsymbol{\Sigma}_e^{-1}, \boldsymbol{y})$ and $P(\boldsymbol{\beta}|\boldsymbol{\Sigma}_e^{-1}, \boldsymbol{y})$ are in the multivariate normal form of order λ.

Appendix B

TABLE B1 *Coefficients $\beta_{j\tau}$ and δ_{jw}—DRG 138: Cardiac Arrhythmia and Conduction Disorders with CC*

Variables	Mean	Standard deviation	Credible intervals boundaries (5°–95°)	
Mental outcome				
Intercept	12.9944	9.7108	7.9759	18.0591
Age	−0.3677	0.4267	−0.5846	−0.1444
Gender	−1.1749	1.0093	−1.6977	−0.6655
Length of stay	0.1200	0.1143	0.0626	0.1782
Initial status of the patient	−0.6052	0.0507	−0.6315	−0.5792
1st readmission	1.5830	1.4000	0.8707	2.3083
2nd readmission	−2.4185	1.3930	−3.1504	−1.6834
Case-mix	12.4720	8.4112	8.0994	16.7934
Revenue / costs ratio	4.0766	2.3789	2.8512	5.3178
Number of beds	−11.8361	12.7455	18.4425	−5.5421
Number of doctors	−13.5155	7.8799	17.4642	−9.4672
Number of nursing personnel	22.1272	16.6445	13.6051	30.6515
Number of operating rooms	−1.2530	5.2501	−4.0106	1.4976
Physical outcome				
Intercept	53.0006	9.2607	48.2590	57.6596
Age	−2.9215	0.4597	−3.1556	−2.6847
Gender	−4.9912	1.0909	−5.5494	−4.4275
Length of stay	0.3115	0.1218	0.2486	0.3748
Initial status of the patient	−0.6460	0.0480	−0.6707	−0.6216
1st readmission	−2.6582	1.4573	−3.4251	−1.8871
2nd readmission	0.5799	1.4285	−0.1812	1.3356
Case-mix	−31.7312	8.5062	−36.1835	27.3803
Revenue / costs ratio	−2.1048	2.5044	−3.4052	−0.8335
Number of beds	9.5460	13.3204	2.6505	16.6378
Number of doctors	−23.8880	17.6963	−32.9683	15.0675
Number of nursing personnel	20.4335	8.1366	16.3149	24.6361
Number of operating rooms	−13.2090	5.5442	−16.1118	10.2839

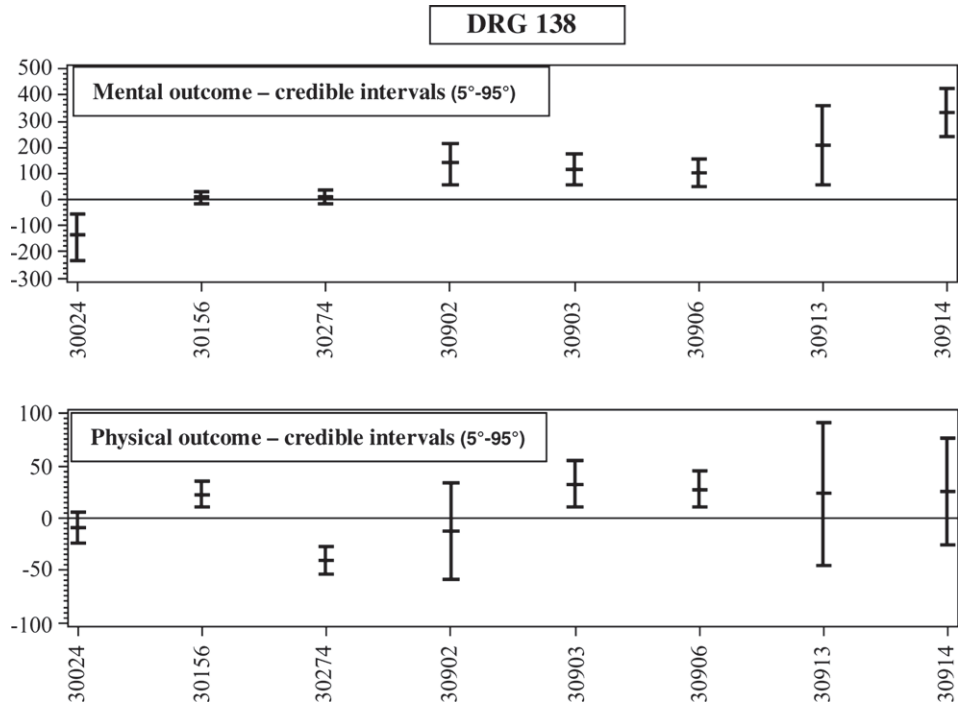

FIG. B2. Credible intervals for m_{jv} (relative effectiveness)—DRG 138: Cardiac Arrhythmia and Conduction Disorders with CC.

TABLE B2 *Matches and ranking of hospitals—DRG 138: Cardiac Arrhythmia and Conduction Disorders with CC*

	30024	30156	30274	30902	30903	30906	30913	30914	Total	Rank
Mental outcome										
30024		−1	1	0	−1	−1	0	0	−2	8°
30156	1		1	0	0	0	0	0	2	1°
30274	−1	−1		0	1	1	0	−1	−1	7°
30902	0	0	0		0	0	0	0	0	3°
30903	1	0	−1	0		0	0	0	0	3°
30906	1	0	−1	0	0		0	0	0	3°
30913	0	0	0	0	0	0		0	0	3°
30914	0	0	1	0	0	0	0		1	2°
Physical outcome										
30024		−1	−1	−1	−1	−1	−1	−1	−7	8°
30156	1		0	−1	−1	−1	−1	−1	−4	7°
30274	1	0		0	−1	−1	−1	−1	−3	6°
30902	1	1	0		0	0	0	−1	1	5°
30903	1	1	1	0		0	0	−1	2	3°
30906	1	1	1	0	0		0	−1	2	3°
30913	1	1	1	0	0	0		0	3	2°
30914	1	1	1	1	1	1	0		6	1°

TABLE B3 *Coefficients $\beta_{j\tau}$ and δ_{jw}—DRG 243: Medical Back Problems*

Variables	Mean	Standard deviation	Credible intervals boundaries (5°–95°)	
Mental outcome				
Intercept	33.0239	20.9216	22.0694	43.9973
Age	−1.6437	0.3276	−1.8183	−1.4786
Gender	−1.7255	0.6822	−2.0860	−1.3744
Length of stay	0.1248	0.0581	0.0948	0.1542
Initial status of the patient				
1st readmission	−2.4644	0.4598	−2.7037	−2.2245
2nd readmission	2.0356	0.6180	1.7165	2.3510
Case-mix	21.1294	10.5171	15.7609	26.5980
Revenue/costs ratio	−20.3868	11.7200	−26.5317	−14.4540
Number of beds	−10.6142	40.5823	−32.1399	9.9538
Number of doctors	−11.8827	9.8347	−16.9848	−6.6896
Number of nursing personnel	16.8247	37.2695	−2.9458	36.2514
Number of operating rooms	1.9809	13.5498	−5.0880	9.0541
Physical outcome				
Intercept	−15.0200	22.6890	−26.9133	−3.5256
Age	−1.2539	0.3569	−1.4401	−1.0731
Gender	−2.1114	0.7643	−2.4970	−1.7153
Length of stay	−0.0882	0.0648	−0.1221	−0.0547
Initial status of the patient	−0.6652	0.0300	−0.6808	−0.6502
1st readmission	−1.2348	0.5109	−1.4912	−0.9724
2nd readmission	0.1122	0.6566	−0.2266	0.4437
Case-mix	−33.1594	11.2902	−39.0753	−27.3606
Revenue/costs ratio	46.5248	12.8400	39.9314	53.0148
Number of beds	−101.1315	44.3381	−123.8517	−78.5247
Number of doctors	16.7576	10.8661	11.1008	22.2761
Number of nursing personnel	109.4583	41.4942	87.7564	130.9669
Number of operating rooms	−39.9162	15.0549	−47.7054	−32.3657

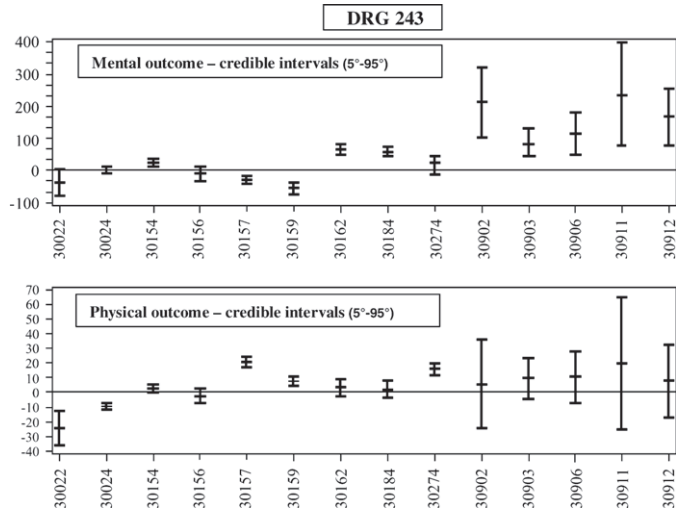

FIG. B3. Credible intervals for m_{jv} (relative effectiveness)—DRG 243: Medical Back Problems.

TABLE B4 *Matches and ranking of hospitals—DRG 243: Medical Back Problems*

	30022	30024	30154	30156	30157	30159	30162	30184	30274	30902	30903	30906	30911	30912	Total	Rank
Mental outcome																
30022		−1	−1	−1	−1	−1	−1	−1	−1	0	−1	−1	0	0	−10	14°
30024	1		−1	0	−1	−1	−1	−1	−1	0	−1	−1	0	0	−7	13°
30154	1	1		0	−1	−1	0	0	−1	0	0	0	0	0	−1	11°
30156	1	0	0		−1	−1	0	0	−1	0	0	0	0	0	−2	12°
30157	1	1	1	1		1	1	1	0	0	0	0	0	0	5	2°
30159	1	1	1	1	−1		0	0	−1	0	0	0	0	0	2	3°
30162	1	1	0	0	−1	0		0	−1	0	0	0	0	0	0	6°
30184	1	1	0	0	1	0	0		−1	0	0	0	0	0	0	6°
30274	1	1	1	1	0	1	1	1		0	0	0	0	0	7	1°
30902	0	0	0	0	0	0	0	0	0		0	0	0	0	0	6°
30903	1	1	0	0	0	0	0	0	0	0		0	0	0	2	3°
30906	1	1	0	0	0	0	0	0	0	0	0		0	0	2	3°
30911	0	0	0	0	0	0	0	0	0	0	0	0		0	0	6°
30912	0	0	0	0	0	0	0	0	0	0	0	0	0		0	6°
Physical outcome																
30022		0	−1	0	0	0	−1	−1	−1	−1	−1	−1	−1	−1	−9	12°
30024	0		0	0	1	1	−1	−1	0	−1	−1	−1	−1	−1	−5	10°
30154	1	0		0	1	1	−1	−1	0	−1	−1	−1	−1	−1	−4	9°
30156	0	0	0		0	1	−1	−1	0	−1	−1	−1	−1	−1	−6	11°
30157	0	−1	−1	0		0	−1	−1	−1	−1	−1	−1	−1	−1	−10	13°
30159	0	−1	−1	−1	0		−1	−1	−1	−1	−1	−1	−1	−1	−11	14°
30162	1	1	1	1	1	1		0	0	−1	0	0	−1	−1	3	6°
30184	1	1	1	1	1	1	0		0	−1	0	0	−1	−1	3	6°
30274	1	0	0	0	1	1	0	0		−1	0	−1	−1	−1	−1	8°
30902	1	1	1	1	1	1	1	1	1		0	0	0	0	9	1°
30903	1	1	1	1	1	1	0	0	0	0		0	0	0	6	5°
30906	1	1	1	1	1	1	0	0	1	0	0		0	0	7	4°
30911	1	1	1	1	1	1	1	1	1	0	0	0		0	9	1°
30912	1	1	1	1	1	1	1	1	1	0	0	0	0		9	1°