Chilean Journal of Statistics
Vol. 3, No. 1, April 2012, 15–29

Bayesian Statistics
Research Paper

A Bayesian random effects model for survival probabilities after acute myocardial infarction

Alessandra Guglielmi[1], Francesca Ieva[1,*], Anna M. Paganoni[1] and Fabrizio Ruggeri[2]

[1]Department of Mathematics, Politecnico di Milano, Milano, Italy
[2]CNR IMATI, Milano, Italy

(Received: 18 March 2011 · Accepted in final form: 14 July 2011)

Abstract

Studies of variations in health care utilization and outcome involve the analysis of multilevel clustered data, considering in particular the estimation of a cluster-specific adjusted response, covariates effect and components of variance. Besides reporting on the extent of observed variations, those studies quantify the role of contributing factors including patients' and providers' characteristics. In addition, they may assess the relationship between health care process and outcomes. In this article we present a case-study, considering a Bayesian hierarchical generalized linear model, to analyze MOMI^2 (Month Monitoring Myocardial Infarction in Milan) data on patients admitted with ST-elevation myocardial infarction diagnosis; both clinical registries and administrative databanks were used to predict survival probabilities. The major contributions of the paper consist in the comparison of the performance of the health care providers, as well as in the assessment of the role of patients' and providers' characteristics on survival outcome. In particular, we obtain posterior estimates of the regression parameters, as well as of the random effects parameters (the grouping factor is the hospital the patients were admitted to), through an MCMC algorithm. The choice of covariates is achieved in a Bayesian fashion as a preliminary step. Some issues about model fitting are discussed through the use of predictive tail probabilities and Bayesian residuals.

Keywords: Bayesian generalized linear mixed models · Bayesian hierarchical models · Health services research· Logistic regression · Multilevel data analysis.

Mathematics Subject Classification: Primary 62F15 · Secondary 62P10 · 62J12.

1. Introduction

Over recent years there has been a growing interest in the use of performance indicators in health care research, since they may measure some aspects of the health care process, clinical outcomes or disease incidence. Several examples, available in clinical literature; see, e.g., Hasday et al. (2002) and Saia et al. (2009), make use of clinical registries to evaluate performances of medical institutions, helping the health governance to plan activities on real epidemiological evidence and needs and to evaluate the performances of structures they manage, providing knowledge about the number of cases, incidence, prevalence and

*Corresponding author. Francesca Ieva. Dipartimento di Matematica Politecnico di Milano, Piazza Leonardo da Vinci 32 I-20133, Milano, Italy. Email: francesca.ieva@mail.polimi.it

ISSN: 0718-7912 (print)/ISSN: 0718-7920 (online)

http://www.soche.cl/chjs

survival concerning a specific disease. As a worthy contribution of this work, both clinical registry and administrative database were used to model in-hospital survival of acute myocardial infarction patients, in order to point out benchmarks to be used in provider profiling process.

The disease we are interested in is the ST-segment elevation acute myocardial infarction (STEMI): it consists of a stenotic plaque detachment, which causes a coronary thrombosis and a sudden critical reduction of blood flow in coronary vessels. STEMI is characterized by a great incidence (650 - 700 events per month have been estimated only in Lombardia region, whose inhabitants are approximately ten millions) and serious mortality (about 8% in Italy), and in fact it is one of the main causes of death all over the world. A case of STEMI can be diagnosed through the electrocardiogram (ECG), observing the elevation of ST segment, and treated by thrombolytic therapy and/or percutaneous transluminal coronary angioplasty (PTCA), which up to now are the most common procedures. The patients in our data set always undergo directly to a PTCA procedure avoiding the thrombolysis, even if the two treatments are not mutually exclusive. Anyway, good results for any of the two treatments can be evaluated by observing first the in-hospital survival of inpatients, and then quantifying the reduction of ST segment elevation one hour after the intervention. Concerning heart attacks, both survival and quantity of myocardial tissues saved from damage strongly depend on time saved during the process; in this work, we focus on the survival outcome. Anyhow, time has indeed a fundamental role in the overall STEMI health care process. By Symptom Onset to Door time we mean the time since symptoms onset up to the arrival at Emergency Room (ER), and Door to Balloon time (DB time) is the time since the arrival at ER up to the surgical practice of PTCA. Clinical literature strongly stresses the connection between in-hospital survival and procedures time, as attested, e.g., in Cannon et al. (2000), Jneid et al. (2008) and MacNamara et al. (2006).

The presence of differences in the outcomes of health care has been documented extensively in recent years. In order to design regulatory interventions by institutions for instance, it is interesting to study the effects of variations in health care utilization on patients outcomes, in particular examining the relationship between process indicators, which define regional or hospital practice patterns, and outcomes measures, such as patients survival or treatment's efficacy. If the analysis of variations concerns in particular the comparison of the performance of health care providers, it is commonly referred to as provider profiling; see Normand et al. (1997) and Racz and Sedransk (2010). The results of profiling analyses often have far-reaching implications. They are used to generate feedback for health care providers, to design educational and regulatory interventions by institutions and government agencies, to design marketing campaigns by hospitals and managed care organizations, and, ultimately, to select health care providers by individuals and managed care groups.

The major aim of this work is to measure the magnitude of the variations of health care providers and to assess the role of contributing factors, including patients' and providers' characteristics on survival outcome in STEMI patients. Data on health care utilization have a "natural" multilevel structure, usually with patients at the lower level and hospitals forming the upper-level clusters. Within this formulation, two main goals are taken into account: one is to provide cluster-specific estimates of a particular response, adjusted for patient's characteristics, while the other one is to derive estimates of covariates effects, such as differences between patients of different gender or between hospitals. Hierarchical regression modelling from a Bayesian perspective provides a framework that can accomplish both these goals. In particular, this article considers a Bayesian generalized linear mixed model (see Zeger and Karim, 1991) to predict the binary survival outcome by means of relevant covariates, taking into account overdispersion induced by the grouping factor.

We illustrate the analysis on a subset of data collected in the MOMI[2] survey on patients admitted with STEMI diagnosis in one of the structures belonging to the Milano Cardiological Network, using a logit model for the survival probability. For this analysis, patients are grouped by the hospital they have been admitted to for their infarction. Assuming a Bayesian hierarchical approach for the hospital factors yields modelling dependence among the random effects parameters, but also using the data set to make inferences on hospitals which do not have patients in the study, borrowing strength across patients, as well as clustering the hospitals. A Markov chain Monte Carlo (MCMC) algorithm is necessary to compute the posterior distributions of parameters and predictive distributions of outcomes, as well as to use other diagnostic tools, such as Bayesian residuals, for goodness-of-fit analysis. The choice of covariates and link functions was suggested first in Ieva and Paganoni (2011), according to frequentist selection procedures and clinical know-how; however, it was confirmed here using Bayesian tools. We found out that killip first, that is an index of the severity of the infarction, and then age, have a sharp negative effect on the survival probability, while the Symptom Onset to Balloon time has a lighter influence on it. An interesting, novel finding is that the resulting variability among hospitals seems not too large, even if we underlined that four hospitals have a more extreme effect on the survival (one has a positive effect, while the remaining three have a negative effect) then the others. Such finding can be explained by the relative homogeneity among the hospitals, all located in Milano, the region capital. Larger heterogeneity is expected in future when extending the analysis to all the hospitals in the region. The advantages of a Bayesian approach to this problem are more than one: providers' profiling or patients' classification are allowed to be guided not only by statistical but clinical knowledge also, hospitals with low exposure can be automatically included in the analysis, and providers' profiling can be simply achieved through the posterior distribution of the hospital-effects parameters.

To the best of our knowledge, this study is the first example of the use of Bayesian methods in provider profiling using data which arise from the linkage between Italian administrative databanks and clinical registries. This paper shares the same framework of hierarchical generalized linear mixed models as in Daniels and Gatsonis (1999), who examined differences in the utilization of coronary artery bypass graft surgery for elderly heart attack patients treated in hospitals.

The paper is organized as follows. Section 2 illustrates the data set about STEMI in Milano Cardiological Network, while Section 3 describes the main features of the proposed model, with a short discussion on covariates selection. Section 4 and 5 discuss prior elicitation and Bayesian inferences, respectively. Finally, Section 6 presents results of the inference on quantities of interest with a discussion. Some final remarks are reported in Section 7. All the analyses have been performed with WinBUGS; see Lunn et al. (2000) and also `http://www.mrc-bsu.cam.ac.uk/bugs` and R (2009) (version 2.10.1) programs.

2. The STEMI Data Set

A net connecting the territory to hospitals, by a centralized coordination of the emergency resources, has been activated in the Milano urban area since 2001. The aim of a monitoring project on it is the activation of a registry on STEMI to collect process indicators (Symptom Onset to Door time, first ECG time, Door to Balloon time and so on), in order to identify and develop new diagnostic, therapeutic and organizational strategies to be applied to patients affected by STEMI by Lombardia region, hospitals and 118 organization (the national toll-free number for medical emergencies). To reach this goal, it is necessary to understand which organizational aspects can be considered as predictive of time to treatment reduction. In fact, organizational policies in STEMI health care process concern both 118 organization and hospitals, since a subject affected by an infarction can reach the hospital by himself or can be taken to the hospital by 118 rescue units.

So, in order to monitor the Milano Cardiological Network activity, times to treatment and clinical outcomes, the data collection MOMI^2 was planned and made on STEMI patients, during six periods corresponding to monthly/bimonthly collections. For these units, information concerning mode of admission (on his/her own or by three different types of 118 rescue units), demographic features (sex, age), clinical appearance (presenting symptoms and Killip class at admittance), received therapy (thrombolysis, PTCA), Symptom Onset to Door time, in-hospital times (first ECG time, DB time), hospital organization (for example, admission during on/off hours) and clinical outcome (in-hospital survival) have been listed and studied. The Killip classification is a system used with acute myocardial infarction patients, in order to stratify them in four risk severity classes. Individuals with a low Killip class are less likely to die within the first 30 days after their myocardial infarction than individuals with a high Killip class. The whole MOMI^2 survey consists of 840 statistical units, but in this work we only focus on patients who underwent primary PTCA and belonging to the third and fourth collections, since they are of better quality. Among the resulting PTCA-patients, we selected those who had their own hospital admission registered also in the Public Health Database of Lombardia region, in order to confirm the reliability of the information collected in the MOMI^2 registry. Finally, the considered data set consists of 240 patients.

Previous frequentist analyses on MOMI^2 survey (see Grieco et al., 2008; Ieva, 2008; Ieva and Paganoni, 2010) pointed out that age, total ischemic time (Symptom Onset to Balloon time, denoted by OB) in the logarithmic scale and killip of the patient, are the most significant factors in order to explain survival probability from a statistical and clinical point of view. Here killip is a binary variable, corresponding to 0 for less severe (Killip class equal to 1 or 2) and 1 for more severe (Killip class equal to 3 or 4) infarction. This choice of covariates was confirmed using Bayesian variable selection procedure; see the next section for more details.

The main goal of our study is to explain and predict, by means of a Bayesian random effects model, the in-hospital survival (i.e., the proportion of patients discharged alive from the hospital). The data set consists of $n = 240$ patients who were admitted to $J = 17$ hospitals after a STEMI event. The number of STEMI patients per hospital ranges from 1 to 32, with a mean of 14.12. Each observation y_i records if a patient survived after STEMI, i.e., $y_i = 1$ if the ith patient survived, $y_i = 0$ otherwise. In the rest of the paper, $\mathbf{y}$ denotes the vector of all responses $(y_1, \ldots, y_n)$. The data set is strongly unbalanced, since 95% of the patients have been discharged alive. The observed hospital-survival rates ranges from 75% to 100%. These high values are explained because they are in-hospital survival probabilities, a follow-up data being not available yet. The data set contained some missing covariates, with proportions of 7%, 24% and 2% for age, OB and killip respectively. The missing data for age and OB were imputed as the empirical means (64 years for age, 553 minutes for OB), while we sampled the missing 0-1 killip class covariates from the Bernoulli distribution with probability of success estimated from the non-missing data. After having imputed all the covariates, the mean value of age and OB did not change, while the proportion of patients with less severe infarction (killip = 0) was 94%. Finally, we had no missing data concerning hospital of admission and outcome.

3. A Bayesian Generalized Mixed-Effects Model

We considered a generalized mixed-effects model for binary data from a Bayesian viewpoint. For a recent review on this topic, see Chapters 1–3 in Dey et al. (2000). For each patient ($i = 1, \ldots, n$), let Y_i be a Bernoulli random variable with mean p_i, which represents the probability that the ith patient survived after STEMI. The p_i's are modelled through a logit regression with covariates $\mathbf{x} := \{\mathbf{x}_i\}$, $\mathbf{x}_i := (1, x_{i1}, x_{i2}, x_{i3})$ which represent the age,

the Symptom Onset to Balloon time in the log scale (log-OB) and the killip, respectively, of the ith patient in the data set. Moreover, age and log-OB have been centered. Since the patients come from J different hospitals, we assume the following multilevel model, with the hospital as a random effect:

$$Y_i|p_i \overset{\text{ind}}{\sim} \text{Be}(p_i), \qquad i = 1, \ldots, n, \tag{1}$$

and

$$\text{logit}(p_i) = \log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + b_{k[i]}, \tag{2}$$

where $b_{k[i]}$ represents the hospital effect of the ith patient in hospital $k[i]$. We denote by $\boldsymbol{\beta}$ the vector of regression parameters $(\beta_0, \beta_1, \beta_2, \beta_3)$. It is well-known that Equations (1) and (2) have a latent variable representation (see Albert and Chib, 1993), which can be very useful in performing Bayesian inference, as well as in providing medical significance: conditioning on the latent variables $Z_1, \ldots, Z_n$, the $Y_1, \ldots, Y_n$ are independent, and, for $i = 1, \ldots, n$,

$$Y_i = \begin{cases} 1, & \text{if } Z_i \geq 0; \\ 0, & \text{if } Z_i < 0; \end{cases} \tag{3}$$

where

$$Z_i = \mathbf{x}_i^\top \boldsymbol{\beta} + b_{k[i]} + \varepsilon_i, \quad \varepsilon_i \overset{\text{i.i.d.}}{\sim} f_\varepsilon, \tag{4}$$

being $f_\varepsilon(t) = \mathrm{e}^{-t}(1+\mathrm{e}^{-t})^{-2}$ the standard logistic density function. The same class of models, however, without considering random effects, was applied in Souza and Migon (2004) to a similar data set of patients after acute myocardial infarction.

As mentioned in Section 2, the choice of covariates was first suggested in Ieva and Paganoni (2011), using frequentist model choice tools. However, we have considered it also in a Bayesian framework, using the Gibbs variable selection method by Dellaportas et al. (2002). But first, as a default analysis, we considered covariates selection via the R package BMA; see Raftery et al. (2009). A subgroup of 197 patients with 11 non-missing covariates was processed by the function `bic.glm`, and 7 covariates were selected (age, OB time, killip, sex, admission during on/off hours, ECG time, number of previous hospitalizations). For this choice of covariates, the non-missing data extracted from the 240-patients data set consists of 217 units, which were again analyzed via `bic.glm`. The posterior probability that each variable is non-zero was very high (about 40%) for age and killip, while they were smaller than 7% for the others. Moreover, the smallest BICs denoting the "best" models resulted for those including age, killip and sex. Since sex is strongly correlated with age in our data set (only elderly women are in), at the end, we agreed with the choice of covariates in Ieva and Paganoni (2011), considering only age and killip, while the OB time was strongly recommended by clinical and health care utilization know-how, since it was the main process indicator of the MOMI^2 clinical survey.

As a second analysis, we consider only covariates which have non-missing values for all patients (age, OB time, killip, sex, admission during on/off hours, number of previous hospitalizations), to be analyzed using the Gibbs variable selection method. The linear predictor assumed in the right hand-side of Equation (2) to select covariates can be represented as

$$\eta_i = \beta_0 + \sum_{j=1}^{6} \gamma_j \beta_j x_{ij}, \quad i = 1, \ldots, n, \tag{5}$$

where $(\gamma_1, \ldots, \gamma_6)$ is a vector of parameters in $\{0, 1\}$. Of course, a prior for both the regression parameter $\boldsymbol{\beta}$ and the model index parameter $\boldsymbol{\gamma}$ must be elicited, so that the marginal posterior probability of $\boldsymbol{\gamma}$ suggest which variables must be included in the model. We assumed different "noninformative" priors for the logit model with the linear predictor given in Equation (5), as suggested in Ntzoufras (2002), implementing a simple BUGS code to compute the marginal posterior distributions for each γ_j, for $j = 1, \ldots, 6$, and the posterior inclusion probabilities. However the analysis confirmed the previously selected model.

The selection of such a few number of covariates (with respect to 13, the total number) is not surprising since previous analyses; see Ieva (2008) and Ieva and Paganoni (2010) pointed out that the covariates are highly correlated. For instance, there is dependency between age on one hand and sex, or symptoms, or mode of admission, on the other, between symptoms and killip, or symptoms and mode of admission, and between sex and symptoms. These relationships can be explained because acute coronary syndromes, as STEMI, affect mainly male patients instead of females, and are more frequent as the patient age increases. Moreover, it is well-known that the STEMI symptoms depend on the severity of the infarction itself, and elderly patients have usually more atypical symptoms. Furthermore, the symptoms may influence the choice of the type of ambulance sent to rescue the patient; ambulances which allow the ECG teletransmission are usually sent to patients presenting more typical infarction symptoms, in order to allow them to skip the waiting time due to ER procedures, and to reduce accordingly the door to balloon time.

4. The Prior Distribution

As mentioned in the previous sections, one of the aim of this paper is to make a comparison among the patients survival probabilities treated in different hospitals of the Milano Cardiological Network. Such an aim can be accomplished if, for instance, we assume the hospital each patient was admitted to as a random factor. We make the usual (from a Bayesian viewpoint) random effects assumption for the hospitals, that is, the hospital effect parameters b_j's are drawn from a common distribution; moreover, since no information is available at the moment to distinguish among the hospitals, we assume symmetry among the hospital parameters themselves, i.e., $b_1, \ldots, b_J$ can be considered as (the first part of an infinite sequence of) exchangeable random variables. Via Bayesian hierarchical models, not only we model dependence among the random effects parameters $\mathbf{b} := (b_1, \ldots, b_J)$, but it be also possible to use the data set to make inferences on hospitals which have few or no patients in the study, borrowing strength across hospitals. As usual in the hierarchical Bayesian approach, the regression parameter $\boldsymbol{\beta}$ and the hospital parameter $\mathbf{b}$ are assumed a priori independent, $\boldsymbol{\beta}$ is given a (multivariate) Gaussian distribution and $\mathbf{b}$ is given a scale-mixture of (multivariate) Gaussian distributions; more specifically:

$$\begin{aligned} &\boldsymbol{\beta} \perp \mathbf{b}, \quad \boldsymbol{\beta} \sim \mathrm{MN}(\boldsymbol{\mu}_\beta, V_\beta), \\ &b_1, \ldots, b_J | \sigma \overset{\text{i.i.d.}}{\sim} \mathrm{N}(\mu_b, \sigma^2), \quad \text{and} \quad \sigma \sim \mathrm{U}(0, \sigma_0). \end{aligned} \tag{6}$$

Observe that the prior assumption on $\mathbf{b}$ is that, conditionally on the parameter σ, each hospital effect parameter has a Gaussian distribution with variance σ^2; here the uniform prior on σ is set as an assumption of ignorance/symmetry on the standard deviation of each hospital effect. The Gaussian prior for $\boldsymbol{\beta}$ is standard, but its hyperparameters, as well as the hyperparameter of the prior distribution for σ, it is given informatively, using available information from other MOMI^2 collections; for more details, see Section 6.2. On the other hand, a more standard prior for b_j would be a scale-mixture of normals, mixed

by an inverse-gamma distribution for σ^2, with parameter (η, η) for small η. However, this prior has been often criticized (see Gelman, 2006), mainly because the inferences do not result robust with respect to the choice of η, and the prior density (for all small η), as well as the resulting posterior, are too peculiar. In what follows, the parameter vector $(\boldsymbol{\beta}, \mathbf{b}, \sigma)$ is denoted by θ.

5. Bayesian Inference

Based on given priors and likelihood, the posterior distribution of θ is expressed by

$$\begin{aligned}\pi\left(\theta|\mathbf{y},\mathbf{x}\right) &\propto \pi\left(\theta\right)\mathcal{L}(\mathbf{y}|\theta,\mathbf{z},\mathbf{x})f(\mathbf{z})\\ &= \pi(\boldsymbol{\beta})\pi(\mathbf{b}|\sigma)\pi(\sigma)\prod_{i=1}^{n}(\mathbb{I}_{(0,+\infty)}(z_i))^{y_i}(\mathbb{I}_{(-\infty,0]}(z_i))^{1-y_i}\prod_{i=1}^{n}f_{\varepsilon}(z_i-\mathbf{x}_i^{\top}\boldsymbol{\beta}-b_{k[i]}).\end{aligned} \tag{7}$$

We are interested in predictions too. This implies (i) considering the posterior predictive survival probability of a new patient coming from an hospital already included in the study, or (ii) the posterior predictive survival probability of a new patient coming from a new $(J+1)$th hospital. We have

$$\mathrm{P}(Y_{n+1}=1|\mathbf{y},\mathbf{x},b_j)=\int_{\mathbb{R}^4}\mathrm{P}(Y_{n+1}=1|\boldsymbol{\beta},b_j,\mathbf{x})\pi(\boldsymbol{\beta}|b_j,\mathbf{y})d\boldsymbol{\beta},\ j=1,\ldots,J, \tag{8}$$

for a new patient with covariate vector $\mathbf{x}$ coming from the jth hospital in the study, and

$$\mathrm{P}(Y_{n+1}=1|\mathbf{y},\mathbf{x},b_{J+1})=\int_{\mathbb{R}^4}\mathrm{P}(Y_{n+1}=1|\boldsymbol{\beta},b_{J+1},\mathbf{x})\pi(\boldsymbol{\beta}|b_{J+1},\mathbf{y})\,d\boldsymbol{\beta}, \tag{9}$$

where $\pi(\boldsymbol{\beta}|b_{J+1},\mathbf{y})$ is computed from

$$\pi(\boldsymbol{\beta},b_{J+1}|\mathbf{y})=\int_{\mathbb{R}^+}\pi(b_{J+1}|\sigma)\pi(\boldsymbol{\beta},\sigma|\mathbf{y})d\,\sigma,$$

being $\pi(b_{J+1}|\sigma)$ the prior population conditional distribution given in Equation (6).

As far as model checking is concerned, we consider predictive distributions for patients already enrolled in the study in the spirit of replicated data in Gelman et al. (2004). More specifically, we compute

$$\mathrm{P}(Y_i^{\mathrm{new}}=1|\mathbf{y},\mathbf{x_i},b_{k[i]}),\quad \text{for all } i=1,\ldots,n. \tag{10}$$

Here, Y_i^{new} denotes the ith "replicated data that could have been observed, or, to think predictively, as the data that we would see tomorrow if the experiment that produced y_i today were replicated with the same model and the same value of parameters that produced the observed data"; see Gelman et al. (2004, Section 6.3). Since we have a very unbalanced data set, the following Bayesian rule is adopted: a patient is classified as alive if $\mathrm{P}(Y_i^{\mathrm{new}}=1|\mathbf{y},\mathbf{x_i},b_{k[i]})=\mathbb{E}[Y_i^{\mathrm{new}}|\mathbf{y},\mathbf{x_i},b_{k[i]}]$ is greater than the empirical mean $\bar{y}_n$. This

rule is equivalent to minimize the expected value of the following loss function

$$L(\mathrm{P}(Y_i = 1|\mathbf{y}, \mathbf{x_i}, b_{k[i]}), a_1) = \mathrm{Max}\{0, \bar{y}_n - \mathrm{P}(Y_i = 1|\mathbf{y}, \mathbf{x_i}, b_{k[i]})\},$$
$$L(\mathrm{P}(Y_i = 1|\mathbf{y}, \mathbf{x_i}, b_{k[i]}), a_0) = \mathrm{Max}\{0, \mathrm{P}(Y_i = 1|\mathbf{y}, \mathbf{x_i}, b_{k[i]}) - \bar{y}_n\},$$

where the action a_1 is to classify the patient as alive and the action a_0 corresponds to classify the patient as dead. Then the coherence between the Bayesian rule and the data set is checked.

Finally we computed the latent Bayesian residuals for binary data as suggested in Albert and Chib (1995). Thanks to the latent variable representation in Equations (3) and (4) of the model, we can consider the realized errors

$$e_i = Z_i - (\mathbf{x}_i^\top \boldsymbol{\beta} + b_{k[i]}), \quad i = 1, \ldots, n, \tag{11}$$

obtained solving Equation (4) w.r.t. ε_i. Each e_i is a function of the unknown parameters, so that its posterior distribution can be computed through the MCMC simulated values, and later examined for indications of possible departures from the assumed model and the presence of outliers; see also Chaloner and Brant (1998). Therefore, it is sensible to plot credibility intervals for the marginal posterior of each e_i, comparing them to the marginal prior credibility intervals (of the same level).

6. Data Analysis

In this section we illustrate the Bayesian analysis of the data set described in Section 2, giving some details on computations and prior elicitation.

6.1 Bayesian computations

As we mentioned in Section 1, all estimates were derived using WinBUGS. The computation of the full conditionals to directly implement a Gibbs sampler algorithm can be computed starting from Equation (7); however they are not "standard" distributions, i.e., closed form expressions do not exist for all of them, given the priors in Equation (6). Some details on the full conditionals for general design GLMMs required by WinBUGS are in Zhao et al. (2006).

The first 100,000 iterations of the chain were discarded, retaining parameter values each 80 iterations to decrease autocorrelations, with a final sample size equal to 5,000; we run the chains much longer (for a final sample size of 10,000 iterations), but the gain in the MC errors was relatively small. Some convergence diagnostics (Geweke's and the two Heidelberger-Welch ones) were checked; see, e.g., the reference manual of the CODA package (Plummer et al., 2006) for more details. Moreover, we monitored traceplots, autocorrelations and MC error/posterior standard deviation ratios for all the parameters, indicating the MCMC algorithm converged. Code is available from the authors upon request.

6.2 Informative prior hyperparameters

Concerning information about hyperprior parameters, we fixed $\mu_b = 0$ regardless of any information, since, by the exchangeability assumption, the different hospitals have the same prior mean (fixed equal to 0 to avoid confounding with β_0). As far as $\boldsymbol{\beta}$ is concerned, we have enough past data to be relatively informative in eliciting prior hyperparameters; they

were fixed after having fitted model given in Equations (1) and (2), under non-informative priors for θ, to "similar" data, i.e., 359 patients undergone primary PTCA whose data were collected during the other four MOMI^2 collections. Therefore, for the present analysis, we fixed $\boldsymbol{\mu}_\beta = (3, 0, 0.1, -0.7)^\top$, which are the posterior means of the regression parameters under the preliminary analysis. The matrix V_β was assumed diagonal, $V_\beta = \text{diag}(2, 0.04, 0.5882, 3.3333)$, which, except for the second value, are about 10 times the posterior variances of the regression parameters under the preliminary analysis (0.04 is 100 times the posterior variance, in order to consider a vaguer prior for β_1). The prior hyperparameter σ_0 was fixed equal to 10, a value compatible with the support of the posterior distribution for σ in the preliminary analysis. Posterior estimates of $\boldsymbol{\beta}$, $\mathbf{b}$ and σ proved to be robust with respect to $\boldsymbol{\mu}_\beta$ and V, even when we fixed a non-diagonal matrix for V, assuming prior dependence through the regression parameters (the non-diagonal V elicited via the preliminary analysis as well). As far as the variances of the β's parameters are concerned, the robustness analysis pointed out that assuming smaller values than those reported here yielded a "too informative" prior, that is the data did not swamp the prior; on the other hand, larger variances produced typical computational difficulties of a "too vague" prior. This choice of the variances values represents an optimal trade-off between these two behaviors.

6.3 Results

Summary inferences about regression parameters and σ can be found in Table 1, while the marginal posterior distributions are depicted in Figures 1 and 2.

Table 1. Posterior means, standard deviations, and 95% credibility intervals of the regression parameters and σ.

		Informative prior		Credibility intervals	
		mean	sd	lwr	upr
intercept	β_0	3.8160	0.5704	2.8310	5.1100
age	β_1	-0.0792	0.0324	-0.1464	-0.0183
log(OB)	β_2	-0.1527	0.3326	-0.7902	0.5154
killip	β_3	-1.5090	0.8159	-3.0470	0.1340
random effect std. dev.	σ	1.1770	0.7417	0.0766	2.8960

From Table 1 and Figure 1 it is clear that the marginal posteriors of β_1 and β_3 are concentrated on the negative numbers, confirming the naïve interpretation that an increase in age or killip class decreases the survival probability. The negative effect of the log(OB) is questionable, given its high variability, even if the posterior median of β_2 is -0.16. Anyway, it was indeed included because of its clinical relevance; moreover, it is the main process indicator in health care monitoring of STEMI procedures. Observe that the posterior mean of $\beta_0 + b_j$, which is the logit of the survival probability for a patient with "average" covariates from any hospital, is between 2.90 and 4.78, yielding a high posterior estimates of the survival probability from any hospital, as expected.

By inspecting Figure 2 a shrinkage of the posterior density of σ with respect to the uniform prior can be observed; this fact supports the conjecture of a low variability within medical institutions, which can be partly explained by the relative homogeneity among the hospitals, all located in Milano. As far as the marginal posterior distribution of the random effect parameters are concerned, Figure 3 displays the posterior median and mean (with 95% credibility intervals) of each hospital parameter b_j, for $j = 1, \ldots, J$.

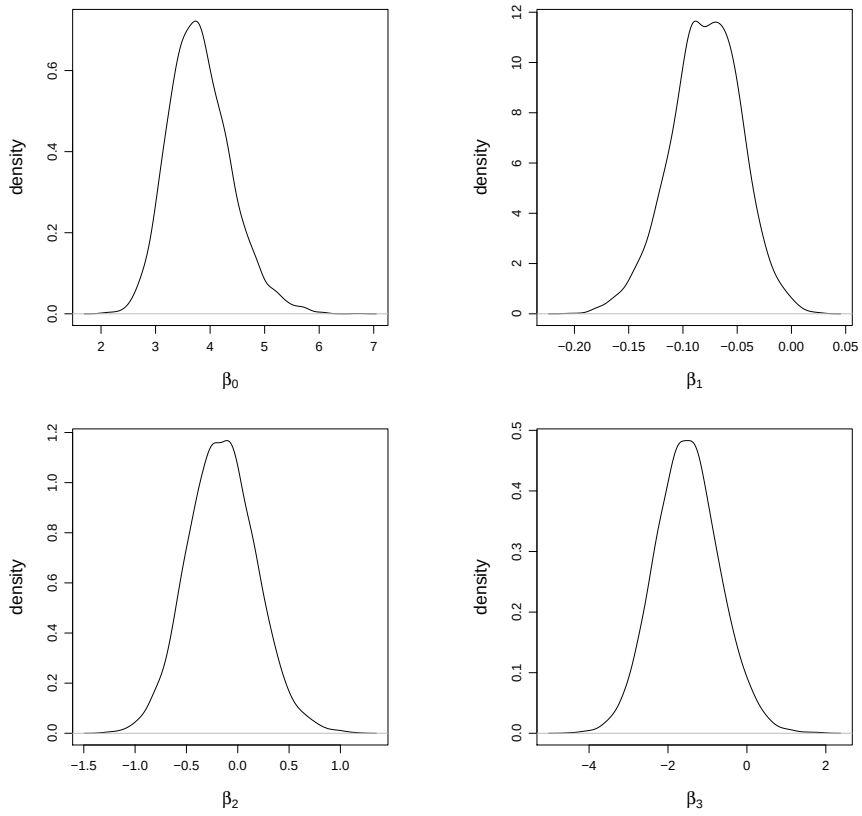

Figure 1. Marginal posterior density of the regression coefficients.

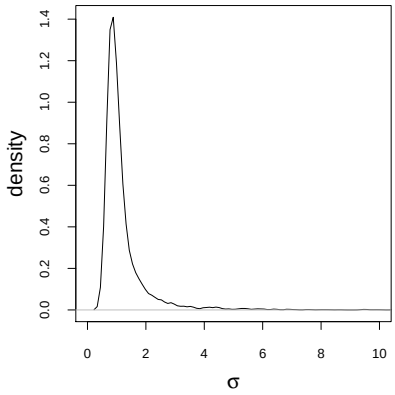

Figure 2. Marginal posterior density of σ.

In Table 2, we report

$$\tilde{p}_j = \min\{\mathrm{P}(b_j > 0|\mathbf{y}), \mathrm{P}(b_j < 0|\mathbf{y})\}, \quad j = 1, \ldots, J,$$

together with the signum of the posterior median of the b_j's. Low values of $\tilde{p}_j$ denote the posterior distribution of b_j is far from 0, so that the jth hospital significantly contributes to the (estimated) regression intercept $\beta_0 + b_j$. In Figure 3, the credible intervals corresponding to $\tilde{p}_j$ less than 0.18 are depicted in yellow; it is clear that hospital 9 has a positive effect, while hospital 10, 11 and 15 have a negative effect on the survival probability.

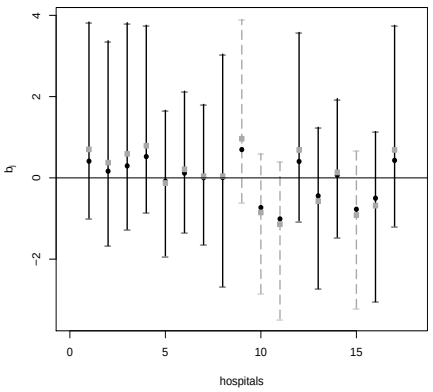

Figure 3. Posterior median (bullet), mean (square) and 95% credibility intervals of all random effect parameters b_j. The credible intervals for hospitals such that $\min(\mathrm{P}(b_j > 0|\mathbf{y}), \mathrm{P}(b_j < 0|\mathbf{y})) < 0.18$ are dashed.

Table 2. Values of $\tilde{p}_j$ and the signum of the posterior median of each hospital parameters.

b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9
0.27	0.40	0.32	0.25	0.44	0.41	0.49	0.49	0.18
+	+	+	+	-	+	+	+	+
b_{10}	b_{11}	b_{12}	b_{13}	b_{14}	b_{15}	b_{16}	b_{17}	
0.17	0.12	0.28	0.28	0.44	0.17	0.26	0.29	
-	+	-	-	+	-	-	+	

Observe that all the credible intervals of the random effect parameters in Figure 3 include 0, so that we might wonder if the random intercept should be discarded from the model. However, Mauri (2011) presents a Bayesian selection analysis of the same data set considered here, concluding that the posterior inclusion probability of the random effect is significantly larger than 0 (between 0.2 and 0.6 under different reasonable priors). Similar findings were drawn in Ieva and Paganoni (2010) from a frequentist perspective.

Figure 4 displays medians and 95% credibility intervals for the posterior predictive survival probabilities give in Equation (8) of four benchmark patients:

(a) $x_1 = 0$, $x_2 = 0$, $x_3 = 0$, i.e., a patient with average age (64 years), average OB (553 min.) and less severe infarction (Killip class 1 or 2);
(b) $x_1 = 0$, $x_2 = 0$, $x_3 = 1$, i.e., a patient with same age and OB as (a), but with severe infarction (Killip class 3 or 4);
(c) $x_1 = 16$, $x_2 = 0$, $x_3 = 0$, i.e., an elder patient (80 years), with average OB (553 min.) and less severe infarction;
(d) $x_1 = 16$, $x_2 = 0$, $x_3 = 1$, i.e., an elder patient with average OB and severe infarction,

coming from an hospital already in the study. The last credibility interval (in red in each panel) corresponds to the posterior predictive survival probability give in Equation (9) of a benchmark patient coming from a new random $(J + 1)$th hospital. Moreover, from the figure it is clear that killip has a stronger (on average) influence on survival than age since, moving from left to right panels (same age, killip increased) the credibility intervals get much wider than moving from the top to the bottom panels (same killip, age increased).

Finally, as far as predictive model checking is concerned, we computed the predictive probabilities in Equation (10); the classification rule described in Section 5 gives an error rate equal to 27% (64 patients were erroneously classified as dead and only 1 patient was

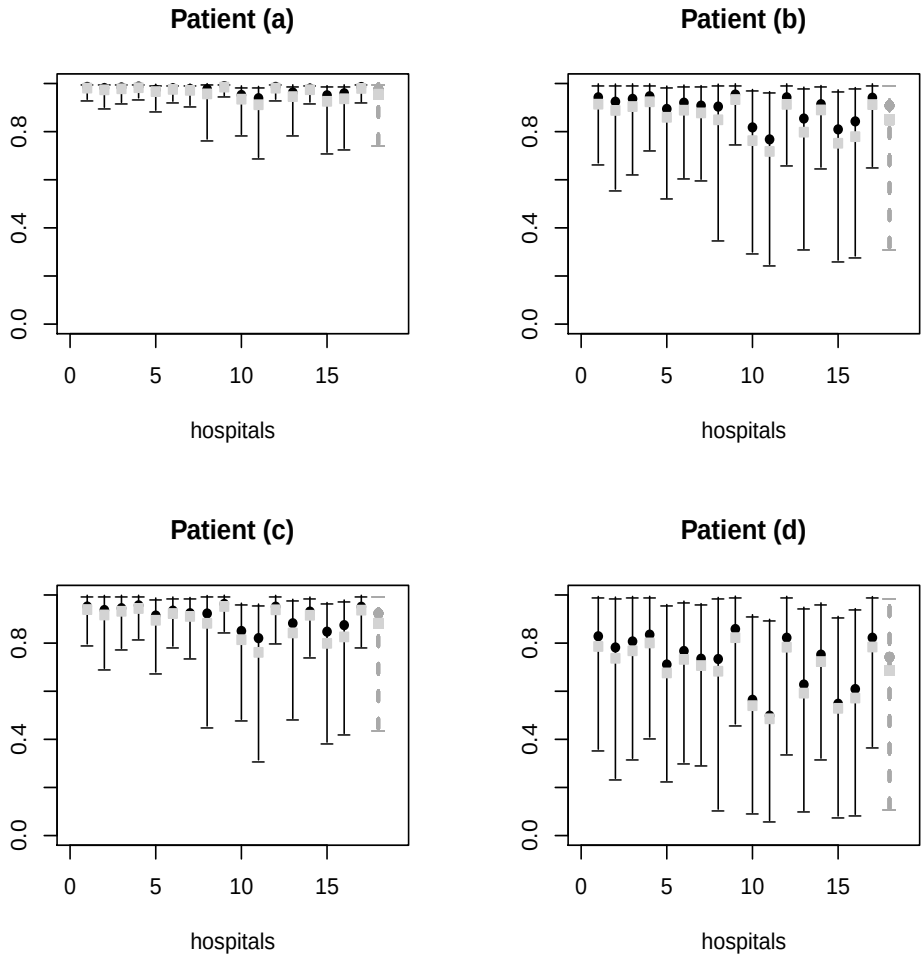

Figure 4. Posterior median (bullet), mean (square) and 95% credible intervals of the posterior predictive survival probabilities for 4 benchmark patients from each hospital in the study and from a new random hospital (the 18th dashed credible interval).

erroneously classified as alive). As a measure of goodness of fit we also computed the Brier score, the average squared deviation between predicted probabilities and outcomes, which is equal to 0.04, showing a fairly good predictive fit of our model.

The left panel of Figure 5 displays the posterior distributions of the Bayesian residuals, as in Equation (11), for each observations, where the red line in the plot denotes the prior marginal distribution (logistic). On the other hand, the right panel shows the same posterior distributions in a 3-dimensional perspective, each residual posterior referring to the posterior survival probability of the corresponding patient.

The picture shows that there are no outliers among the patients who survived, since their posterior residual densities and the prior residual one share the same cluster. More variability appears among the dead patients as far as posterior location and dispersion are concerned. This feature could be brought about by the disparity in the number of cases among the dead and the alive in our data set. Moreover, most deaths occur in the class of more severe infarction, and concern elder people. This rationale explains the larger credibility intervals in Figure 4(d) (right bottom panel) as well, which in fact refers to elderly patients with severe infarction.

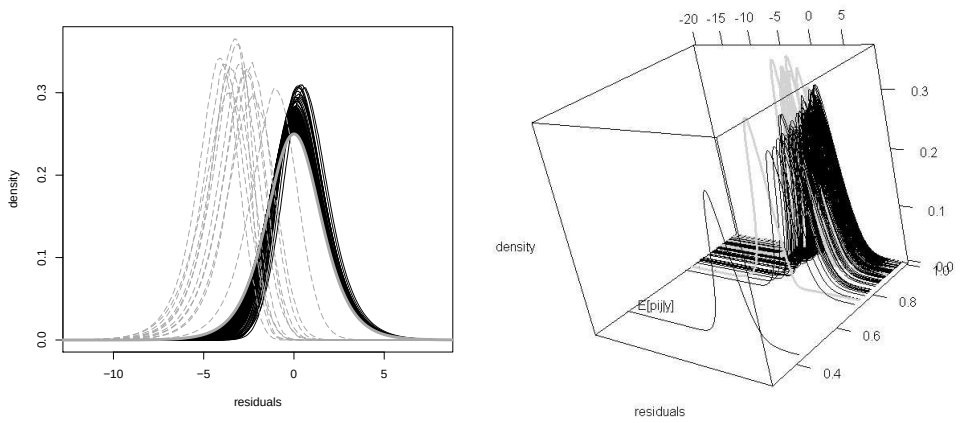

Figure 5. Left panel: posterior distributions of the latent Bayesian residuals. The dashed and solid lines correspond to observations $y_i = 0$ (dead) and $y_i = 1$ (alive), respectively. The solid gray line is the marginal prior distribution (logistic). Right panel: posterior distributions of the latent Bayesian residuals against the expected posterior survival probabilities.

7. Conclusions

In this work we have considered a Bayesian hierarchical generalized linear model with random effects for the analysis of clinical and administrative data with a multilevel structure. These data arise from MOMI^2 clinical registry, based on a survey on patients admitted with ST-elevation myocardial infarction diagnosis, integrated with administrative databanks. The analysis carried out on them could provide a decisional support to the cardiovascular health care governance. We adopted a Bayesian point of view to tackle the problem of modelling survival outcomes by means of relevant covariates, taking into account overdispersion induced by the grouping factor, i.e., the hospital where each patient has been admitted to. To the best of our knowledge, this study is the first example of a Bayesian analysis of data arising from the linkage between Italian administrative databanks and clinical registries. The main aim of this paper was to study the effects of variations in health care utilization on patient outcomes, since the adopted model points out relationships between process and outcome measures. We also provided cluster-specific estimates of survival probabilities, adjusted for patients characteristics, and derived estimates of covariates effects, using MCMC simulation of posterior distributions of the parameters; moreover we discussed model selection and goodness of fit. We found out that Killip first, and age, have a sharp negative effect on the survival probability, while the OB (onset to balloon) time has a lighter influence on it. The resulting variability among hospitals seems not too large, even if we underlined that 4 hospitals have a more extreme effect on the survival: in particular hospital 9 had a positive effect, while hospitals 10, 11 and 15 had a negative effect. As far as negative features of the MCMC outputs are concerned, we found that the marginal posterior distributions of (β_0, b_j), for each j, are concentrated on lines of the whole parameter space, due to the "confounding" between the intercept parameter and the random effects parameters. However the mixing and the convergence of the chain, under a suitable thinning, were completely satisfactory. Finally, as a further step in the analysis, we are considering Bayesian nonparametrics to model the hospital effects, in order to take advantage of the "in-built" clustering they provide.

References

Albert, J.H., Chib, S., 1993. Bayesian analysis of binary and polychotomous response data. Journal of the American Statistical Association, 88, 669–679.

Albert, J.H., Chib S., 1995. Bayesian residual analysis for binary response regression models. Biometrika, 82, 747–759.

Cannon, C.P, Gibson, C.M., Lambrew, C.T., Shoultz, D.A., Levy, D., French, W.J., Gore, J.M., Weaver, W.D., Rogers, W.J., Tiefenbrunn, A.J., 2000. Relationship of symptom-onset-to-balloon time and door-to-balloon time with mortality in patients undergoing angioplasty for acute myocardial infarction. Journal of American Medical Association, 283, 2941–2947.

Chaloner, K., Brant, R., 1988. A Bayesian approach to outlier detection and residual analysis. Biometrika, 31, 651–659.

Daniels, M.J., Gatsonis, C., 1999. Hierarchical generalized linear models in the analysis of variation in health care utilization. Journal of the American Statistical Association, 94, 29–42.

Dellaportas, P., Forster, J.J., Ntzoufras, I., 2002. On Bayesian model and variable selection using MCMC. Statistics and Computing, 12, 27–36.

Dey, D.K., Ghosh, S.K., Mallick, B.M., (eds.) 2000. Generalized Linear Models: A Bayesian Perspective. Chapman & Hall/CRC, Biostatistics Series, New York.

Gelman, A., 2006. Prior distributions for variance parameters in hierarchical models (Comment on article by Browne and Draper). Bayesian Analysis, 3, 515–534.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2004. Bayesian Data Analysis. Second edition. Chapman & Hall/CRC, Boca Raton, Florida.

Grieco, N., Corrada, E., Sesana, G., Fontana, G., Lombardi, F., Ieva, F., Marzegalli, M., Paganoni, A.M., 2008. Predictors of reduction of treatment time for ST-segment elevation myocardial infarction in a complex urban reality. The MoMi2 survey. MOX Report n. 10/2008. Dipartimento di Matematica, Politecnico di Milano. Available at `http://mox.polimi.it/it/progetti/pubblicazioni/quaderni/10-2008.pdf`.

Hasday, D., Behar, S., Wallentin, L., Danchin, N., Gitt, A.K., Boersma, E., Fioretti, P.M., Simoons, M.L., Battler, A., 2002. A prospective survey of the characteristics, treatments and outcomes of patients with acute coronary syndromes in Europe and the mediterranean basin. The euro heart survey of acute coronary syndromes. European Heart Journal, 23, 1190–1210.

Ieva, F., 2008. Modelli statistici per lo studio dei tempi di intervento nell'infarto miocardico acuto. Master Thesis. Dipartimento di Matematica, Politecnico di Milano. Available at `http://mox.polimi.it/it/progetti/pubblicazioni/tesi/ieva.pdf`.

Ieva, F., Paganoni, A.M., 2010. Multilevel models for clinical registers concerning STEMI patients in a complex urban reality: a statistical analysis of MOMI^2 survey. Communications in Applied and Industrial Mathematics, 1, 128–147.

Ieva, F., Paganoni, A.M., 2011. Process indicators for assessing quality of hospitals care: a case study on STEMI patients. JP Journal of Biostatistics, 6, 53–75.

Jneid, H., Fonarow, G., Cannon, C., Palacios, I., Kilic, T., Moukarbel, G.V., Maree, A.O., Liang, L., Newby, L.K., Fletcher, G., Wexler, L., Peterson, E., 2008. Impact of time of presentation on the care and outcomes of acute myocardial infarction. Circulation, 117, 2502–2509.

Lunn, D.J., Thomas, A., Best, N., Spiegelhalter, D., 2000. WinBUGS - a Bayesian modelling framework: concepts, structure, and extensibility. Statistics and Computing, 10, 325–337.

MacNamara, R.L., Wang, Y., Herrin, J., Curtis, J.P., Bradley, E.H., Magid, D.J., Peterson, E.D., Blaney, M., Frederick, P.D., Krumholz, H.M., 2006. Effect of door-to-balloon time on mortality in patients with ST-segment elevation myocardial infarction. Journal of American College of Cardiology, 47, 2180–2186.

Mauri, F., 2011. Bayesian variable selection for logit models with random intercept: application to STEMI data set. Master Thesis. Dipartimento di Matematica, Politecnico di Milano.

Normand, S.T., Glickman, M.E., Gatsonis, C.A., 1997. Statistical methods for profiling providers of medical care: issues and applications. Journal of the American Statistical Association, 92, 803–814.

Ntzoufras, I., 2002. Gibbs variable selection using BUGS. Journal of Statistical Software, 7. Available at `http://www.jstatsoft.org/`.

Plummer, M., Best, N., Cowles, K., Vines, K., 2006. CODA: Convergence diagnosis and output analysis for MCMC. R News, 6, 7–11.

R Development Core Team, 2009. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Available at `http://www.R-project.org`.

Racz, J., Sedransk, J., 2010. Bayesian and frequentist methods for provider profiling using risk-adjusted assessments of medical outcomes. Journal of the American Statistical Association, 105, 48–58.

Raftery, A., Hoeting, J., Volinsky, C., Painter, I., Yeung, K.Y., 2009. BMA: Bayesian model averaging. Available at `http://CRAN.R-project.org/package=BMA`.

Saia, F., Piovaccari, G., Manari, G., Guastaroba, P., Vignali, L., Varani, E., Santarelli, A., Benassi, A., Liso, A., Campo, G., Tondi, S., Tarantino, F., De Palma, R., Marzocchi, A., 2009. Patient selection to enhance the long-term benefit of first generation drug-eluting stents for coronary revascularization procedures: insights from a large multicenter registry. Eurointervention, 5, 57–66.

Souza, A.D.P., Migon, H.S., 2004. Bayesian binary regression model: an application to in-hospital death after AMI prediction. Pesquisa Operacional, 24, 253–267.

Zeger, S.L., Karim, M.R., 1991. Generalized linear models with random effects: a Gibbs sampling approach. Journal of the American Statistical Association, 86, 79–86.

Zhao, Y., Staudenmayer, J., Coull, B.A., Wand, M.P., 2006. General design Bayesian generalized linear mixed models. Statistical Science, 21, 35–51.