

Performance assessment using mixed effects models: a case study on coronary patient care

NICCOLÒ GRIECO

Azienda Ospedaliera Niguarda Cà Granda, 20162 Milano, Italy

AND

FRANCESCA IEVA* AND ANNA MARIA PAGANONI

MOX—Dipartimento di Matematica, Politecnico di Milano, 20133 Milano, Italy

*Corresponding author: francesca.ieva@mail.polimi.it

[Received on 24 August 2010; accepted on 30 March 2011]

Provider profiling is the process of evaluation of the performance of hospitals, doctors and other medical practitioners in order to increase the quality of medical care. This paper reports statistical analyses carried out on data arising from a regular survey concerning patients admitted with an ST-elevation myocardial infarction diagnosis in one of a number of hospitals in the Milan area. The main aim is to determine process indicators to be used in health-care evaluation. Effective statistical support for clinical and organizational governance is obtained by analysing and modelling data from clinical registries. A grouping structure and a consequent ranking of hospitals is investigated, taking into account the fact that this kind of survey data are intrinsically grouped at first level by where patients are hospitalized. We compare three different techniques for hospital classification based, respectively, on: (a) traditional comparison of survival rates; (b) analysis of variance components in fitted generalized linear mixed effects models; and (c) non-parametric random effects estimation.

Keywords: generalized linear mixed models; in-hospital survival; ST-elevation myocardial infarction; provider profiling.

1. Introduction

Performance indicators for assessing quality in health-care contexts have drawn more and more attention over the last few years because they enable the research workers to measure several components of the health-care process, clinical outcomes and disease incidence. At the same time, questions about the right use of such indicators as a measure of care quality have emerged. In this work, we propose the use of performance indicators in modelling the outcomes of clinical structures in order to identify ‘similar behaviours’ among clinical structures. These models include variability between institutions (not forgetting case-mix) and performance indicators are computed starting from data collected through clinical registries. The purpose of this work is, in fact, to highlight how advanced statistical methods can be used to identify suitable models for complex data coming from clinical registries in order to classify and evaluate health-care providers.

In clinical literature (see Saia *et al.*, 2009; Hasday *et al.*, 2002; Dalby *et al.*, 2003), several examples make use of clinical registries to evaluate performances of medical institutions. These databases are very useful because they enable people in charge of the health-care governance to plan activities on real epidemiological evidence and needs. In general, scheduling is strictly connected with a deep knowledge of current health needs, innovative surgery practices acknowledged and measurement of clinical outcomes.

Many accredited procedures for analysis and comparison of the health-care providers are concentrated on the identification of outlier cases. The purpose of our work is different: we have been asked by Regione Lombardia to perform a preliminary analysis in order to capture the real standards of performances of the cardiological Network of Milano, a fairly small area housing nevertheless a lot of providers. Once the real situation has been described, the main interest of Regione Lombardia is to quantify the influence of providers on clinical outcomes and to use information from clinical registries to make the hospital network more effective. This would mean changing funding policies or, at least, signing agreements on opening times in order to rationalize and improve the overall health-care process. Moreover, benchmarks of performance and guidelines based on the real evidence can be pointed out. Then, in this work, we try to identify an effective and robust statistical technique to find similar behaviours or clusters among clinical structures. In general, we could define ‘provider profiling’ as follows: procedures for analysing and comparing the effects of the health-care providers on health services delivery and outcomes. In a typical profiling procedure, patient-level responses are measured by clusters of patients treated by different providers. We believe that statistical analysis of data from clinical registries that are focused on specific diseases might point out suitable indicators of hospitals’ care quality and support decision making in health-care context. Once data have been collected in clinical registries and/or administrative data banks, a study to assess the correlation among hospitals’ care quality and outcomes can be performed.

As we will see in the next sections, in this work, we will use data coming from MOntH MOntoring Myocardial Infarction in MIlan (MOMI)² clinical survey on ST-elevation myocardial infarction (STEMI) and we will model survival outcome by means of suitable patient’s covariates and process indicators, after that we will try to classify hospitals in groups of ‘similar behaviour’. Particularly, we want to propose three different methods to evaluate hospitals performances: in the first, one we estimate the in-hospital survival rates after fitting a generalized linear model (GLM) on outcome of interest, verifying the absence of outliers; in the second, one we fit a generalized linear mixed effects model (GLMM) to explain in-hospital survival outcome, with a parametric random effect due to the hospital as grouping factor; subsequently, we perform an explorative classification analysis implementing a ranking and clustering procedure on the point estimates of hospital effects. Finally, in the third, one we classify hospitals on the basis of the variance components analysis explained by a GLMM on outcome with a non-parametric random effect. Fitting a model with non-parametric random effect enables us to relax normality assumptions; moreover, as it is proved in Lindsay (1983), non-parametric maximum likelihood (NPML) is attained with a discrete measure. The discreteness of random effect induces automatically a natural clustering that highlights similar behaviours in hospitals’ performances.

The article is structured as follows: first, we present from a clinical perspective the cardiovascular diseases, focusing on Regione Lombardia health-care policy (Section 2), then we discuss the role of the statistician and how research domain can influence decision-making context (Section 3). Then, we describe the MOMI² clinical survey (Section 4) a prominent example of innovative use of clinical database as a support in health-care policy in Regione Lombardia, and the statistical models fitted on these data in order to classify hospitals (Section 5). Finally, the results of statistical analysis are presented (Section 6) and discussion of results and suggestions for further developments are proposed (Section 7). All the analyses have been performed with R program (version 2.10.1, R Development Core Team, 2009).

2. Cardiovascular diseases and health policy in Regione Lombardia

It is known that cardiovascular diseases are nowadays one of the main causes of death all over the world. In fact, the American Heart Association reckons that they are the major killers in the clinical context because of the mortality they often induce.

Among them, we are particularly interested in acute coronary syndromes (ACSs) and specifically in acute myocardial infarction (AMI) with STEMI, which is a disease featuring a great incidence (650–700 events per month have been estimated just in Regione Lombardia) and serious mortality (Italy 8%, data coming from ‘Istituto Superiore della Sanità’). As we said before, it is one of the main causes of death all over the world. In general, the AMI is the most frequent disease among the class of ACS. These pathologies are caused by a stenotic plaque detachment, which causes a coronary thrombosis and a sudden critical reduction of blood flow in coronary vessels. This process causes a widespread necrosis of myocardial tissues and leads to an inadequate feeding of myocardial muscle itself.

A case of STEMI can be diagnosed through the electrocardiogram (ECG). It highlights bad patterns, e.g. those showing ST-segment elevation. The subgroup of STEMI patients must be treated as soon as possible (American Heart Association and American College of Cardiology suggest <90 min between arrival at emergency room and treatment time). Up to now, thrombolytic therapy and percutaneous transluminal coronary angioplasty (PTCA) are the most common procedures in dealing with STEMI events. The former consists in a pharmacological treatment that causes a breakdown of the blood clots, whereas in the latter, an empty and deflated balloon driven by a wire, known as *Balloon* catheter, is inserted into the narrowed or obstructed vessels and then inflated to a fixed size. The balloon crushes the fatty deposit, so that the vessel can be opened up, the blood flow improved and finally, the balloon is deflated and withdrawn. In our data (collected in the Milanese urban area), patients always undergo a direct PTCA procedure avoiding the thrombolysis, even if one treatment does not exclude the other. A good practice can be evaluated by observing first the in-hospital survival of inpatients, then quantifying the reduction of ST-segment elevation 1 h after the surgery: if the reduction is over 70%, we could consider the procedure effective. Both the survival and the quantity of myocardial tissue saved from permanent injury depend strongly on the time saved during the process. So, treatment times (i.e. time of first ECG, time existing between arrival at emergency room (ER) and PTCA, . . .) take a key role in influencing outcomes (see Cannon *et al.*, 2000) and stand as candidates of process indicators to evaluate the performance of a clinical institution.

On 11 February 2005, the ‘Piano Cardio-Cerebro Vascolare’ was introduced in Regione Lombardia through D.G.R.20592. This law sets favourable conditions for using clinical registries in health-care process planning. Several clinical registries have been made in Regione Lombardia up to now. In fact, Regione Lombardia is very sensitive to cardiovascular diseases, as it is proved by the huge amount of social and scientific initiatives concerning these syndromes. Among the most important clinical and scientific projects conducted and funded by Regione Lombardia during past years, we can list:

- *Strategic Program* (2008–2011) → Identification and development of new diagnostic, therapeutic and organizational strategies for patients affected by ACSs in order to improve the occurrence of clinical outcomes;
- *Nuove Reti Sanitarie* (2004–2009) → Telemonitoring activities for patients affected by chronic cardiac insufficiency and those concerned with in-home care after cardiac admittance to hospital;
- *PROMETEO—PROgetto Milano Ecg Teletrasmessi ExtraOspedaliero* (2009) → All basic rescue units working in the urban area of Milan have been equipped with ECG telemonitoring machinery;
- *MOMI²—MOnth MOnitoring Myocardial Infarction in Milan* (2006–2008) → A six time-repeated activity of data collection (lasting from 30 to 60 days, MOMI².1–MOMI².6) on STEMI patients in urban area of Milan, performed by Working Group for Cardiac Emergency of Milan, Cardiology Societies and 118 (national free toll number for emergencies) Dispatch Center;

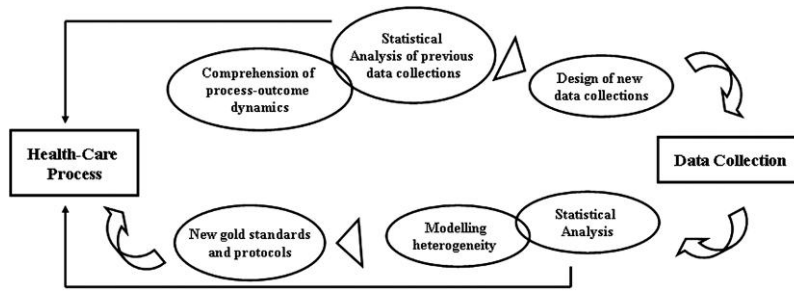


FIG. 1. Flow chart representing the statistician's role in managing the health-care process.

- *GestIMA: Gestione dello STEMI in Lombardia* (2003) → Bimonthly data collection (October 15–November 14th 2003) concerning patients with STEMI diagnosis (see [Oltrona et al., 2005](#)).

It may be argued that control and use of such huge amount of complex data for clinical and epidemiological scopes are very hard tasks. In this context, the role of the statistician and of statistical analysis for management of health-care become central.

In next sections, we will focus our interest and analyses only on MOMI² survey.

3. The role of statistic and statisticians in managing the cardiovascular health care

The role of the statistician in managing the cardiovascular health care is often difficult and controversial. First, he is asked to understand and point out the mechanisms connecting the health-care process itself and the inpatients outcomes; in case of STEMI, they are in-hospital survival and reduction of ST-segment elevation after angioplasty surgery. Within the health-care plan of Regione Lombardia, the statisticians can monitor, analyse and model data collected by clinical surveys or available from already existing databases.

In fact, starting from the existing health-care process,¹ the comprehension of process dynamics depends on statistical analyses of previous data collections. This enable the statistician to give a first feedback to the players involved into the process (hospitals, institutions, clinicians, decision makers and so on) and to plan and realize new data collections for specific needs. The analyses on these new databases enable the statistician to model phenomena and evaluate process indicators in order to point out new gold standards and protocols and to give a further feedback to the involved players, so that health-care process improvements can be obtained. Figure 1 shows the flow chart of the process described above.

As detailed in next sections, in this work, MOMI² survey plays the role of an on-going data collection that the statistician works out in order to give feedback to decision makers and providers involved in health-care process. The main goal of Regione Lombardia is to point out a reliable description of providers conduct to carry out similar behaviours in different hospitals and to quantify the influence of different behaviours on outcomes of interest.

4. The MOMI² clinical survey

The MOMI² project is a result of a collaboration between the Working Group for Cardiac Emergency (ACEU) of Regione Lombardia, Dispatch Center of 118 and Niguarda Cà Granda hospital, concerning

¹In our case, it starts from infarction symptoms onset and ends with pharmaceutical therapy or surgery.

the management of the Network activated in the Milanese urban area since 2001 in order to connect the territory to hospitals by a centralized coordination of the emergency resources. Its primary aims are: (a) promoting the best use of the different reperfusion strategies, (b) reducing transport and decisional delays connected with logistic matters and therapies, (c) increasing the number of patients undergoing primary PTCA before 90 min since the arrival at ER, limit suggested by the American Heart Association/American College of Cardiology guidelines (see *Antman et al., 2008; Ting et al., 2008*). Difficulties in reaching these goals are mainly due to the fact that the Milanese urban area is a complex area with a high density of population (2.9 million resident and 1 million daily commuters) and a high number of hospitals ($n = 27$). Twenty-three of them have a cardiology division and a Critical Care Unit; 18 offer a 24 h available Cath Lab for primary PTCA, 5 are provided with a Cardiac Surgery Unit.

The aim of this project is the activation, on the Milanese urban area, of a registry on AMI to collect process indicators (Symptom Onset time, first ECG time, Door to Balloon time and so on) to be used in profiling providers' service. Specifically, the main purpose of the study is the identification and development of new diagnostic, therapeutic and organizational strategies to be applied (by Regione Lombardia, 118 and hospitals) to patients with STEMI in order to increase the probability of a positive clinical outcome and to improve the health-care offered to patients. In order to do this, it is first necessary to understand which organizational aspects can be considered as predictive of reduction of time to treatment. Therefore, a special attention is focused on the mode of admittance; five different types of patients can be pointed out:

- *self-presented* patients, i.e. patients arriving at the hospital by themselves;
- patients delivered by advanced life support (ALS) units with teletransmission of ECG (ALS + tele-ECG), i.e. rescue units provided with LIFEPAK 12² and doctors onboard;
- patients delivered by ALS, i.e. a rescue unit with doctors onboard but without ECG teletransmission equipment;
- patients delivered by basic life support units, i.e. the common ambulances;
- patients *transferred*, i.e. patients admitted to a certain hospital and then moved to undergo to angioplasty into another one.

Beyond the mode of admittance, further information can be found in the MOMI² data set: for example, personal data like age and sex, clinical data like declared symptoms, Killip class (which quantifies the severity of infarction in four categories) and given therapy, organizational data like mode of admission, hospital and activation of Fast-Track, data concerning all the procedure times and finally, clinical outcomes: in-hospital survival and reperfusion efficacy.

The MOMI² survey is then a retrospective observational study. Anyway, it is a study who enables us to give a 'real-time' feedback on the monitored activities. In fact, the MOMI² survey composed six collections, planned and made during monthly/bimestral periods. In particular:

MOMI ² .1:	90 patients	June 1st–June 30th 2006
MOMI ² .2:	147 patients	November 15th–December 15th 2006
MOMI ² .3:	220 patients	June 1st–July 31st 2007

²A box that enables clinicians to make ECG and send it to the Dispatch Center and to the hospital where the patient will be hospitalized.

MOMI ² .4:	131 patients	November 15th–December 15th 2007
MOMI ² .5:	120 patients	June 1st–June 30th 2008
MOMI ² .6:	133 patients	January 28th–February 28th 2009

The whole data set collects data concerning 841 patients.

In the diagram proposed in Fig. 1, the MOMI² analysis can be seen as the step of analysis of data collection in order to give real-time feedback to providers and to point out critical situations to work out. In fact, the first analysis we have performed (see Ieva, 2008) shows the crucial importance of ECG tele-transmission, and PROMETEO project was planned for equipping all Milan basic rescue units with ECG machinery. Moreover, the results achieved by that work encouraged governance of Regione Lombardia to intensify and widen MOMI² paradigm of data collection and monitoring. The result was *STEMI Archive*, a new data collection on cardiological diseases (see Barbieri *et al.*, 2010; Grieco *et al.*, 2008; Ieva & Paganoni, 2009, 2010 for further details). Finally, the main feedback of our analysis for providers is the hospital classification. This is something we have performed only after modelling outcomes by means of suitable covariates, taking into account process indicators and case mix. In next sections, the use of advanced statistical techniques for clustering providers' behaviour in terms of in-hospital survival will be shown, starting from data of MOMI² survey.

5. Models for classification

The innovative purpose of health-care governance of Regione Lombardia is to use information coming from clinical registries to make the hospital network more effective and then improving the overall health-care process. So three different methodologies to evaluate hospital's performance in this provider profiling perspective are proposed. In the first one, we estimate the in-hospital survival rates after fitting a GLM on outcome of interest and we compute for each hospital a confidence interval for the expected value of in-hospital survivals in order to check the presence of any outlier. In the second one, we fit a GLMM to explain in-hospital survival outcome, with a parametric random effect due to the hospital grouping factor, then we perform an explorative classification and ranking analysis on the point estimates of hospital effects. In the third one, we classify the hospitals on the basis of the variance components analysis explained by a GLMM on outcome with a non-parametric random effect. We will use all the three methods to discriminate between different behaviours. We will compare classification structures obtained starting from these models and we will quantify the effect of making part of different groups on outcomes of interest.

5.1 Generalized linear models

GLMs represent a class of fixed effects regression models for several types of dependent variables (i.e. continuous, dichotomous and counts), see Mc Cullagh & Nelder (2000). Common GLMs include linear regression, logistic regression and Poisson regression.

There are three specifications in a GLM. First, the linear predictor $\eta_i = \mathbf{x}_i' \beta$, where \mathbf{x}_i is the vector of regressors for unit i with fixed effects β . Then, a link function $g(\cdot)$ is specified which converts the expected value μ_i of the outcome variable Y_i (i.e. $\mu_i = \mathbb{E}[Y_i]$) to the linear predictor η_i , i.e. $g(\mu_i) = \eta_i$. Finally, a specification for the form of the variance in terms of the mean μ_i is made. The latter two specifications usually depend on the distribution of the outcome Y_i , which is assumed to belong to the exponential family of distributions. Fixed effects models, which assume that all observations are independent, are not appropriate for analysis of several types of correlated data structures, in particular, for clustered and/or longitudinal data. Anyway they can be considered a first straightforward attempt

to model these data, which are typical in clinical literature. Then, starting from GLM model fitting, we compute the statewide survival rate (SSR) for hospital j , defined as

$$\text{SSR}_j = \frac{\sum_{i=1}^{n_j} y_{ij}^{\text{obs}}}{\sum_{i=1}^{n_j} \hat{p}_{ij}},$$

where y_{ij}^{obs} is the resulting outcome for patient i treated in the hospital j and \hat{p}_{ij} is the corresponding survival probability estimated by using a GLM. SSR_j relates the actual survival at the j -hospital to the expected survival in the same hospital, adjusted for different patient severity resumed in the covariates of the GLM. An elementary assessment of hospital j can be obtained by comparing SSR_j with 1. Moreover, confidence intervals of nominal level 95% for the expected value of $D_j = \sum_{i=1}^{n_j} y_{ij}^{\text{obs}}$ can be computed assuming that D_j follows a Poisson distribution and using Byar's approximation. If the confidence interval for the $E[D_j]$ does not contain $\hat{E}_j = \sum_{i=1}^{n_j} \hat{p}_{ij}$, there is evidence for declaring the hospital as outlier (see Racz & Sedransk, 2010).

5.2 Generalized linear mixed effects models

In clinical contexts, subjects are observed nested within larger units (schools, hospitals, neighbourhoods, workplaces and so on). These are often called 'multilevel' or 'hierarchical' data in which the first level observations (subjects) are nested within the higher second level observations (clusters). Higher levels are also possible: for example, a three-levels model might include repeated observations (first level) nested within subjects (second level) that are nested within clusters (third level).

For statistical analysis of multilevel data, random cluster effects can be added to the regression model in order to take into account the correlation of the data. The resulting model is a mixed model including the usual fixed effects for the regressors plus the random effects. Mixed models for continuous normal outcomes have been extensively developed, anyway many developments have been produced for non-Normal data as well. A lot of these developments fall within GLMMs, which extend GLMs by the inclusion of parametric Gaussian random effect in the linear predictor. An alternative way to assess the performance of hospitals can be obtained by analysing the estimated values of random effect for every hospital.

Since one of the goals of the analyses performed on MOMI² survey is to find a model for grouped data, clustered by providers where patients are hospitalized, our case can be thought as belonging to the class of problems where fixed effects models performance in explaining phenomenon variability is very poor. For this reason, GLMM is straightforward to be considered. In this case, the method for classifying providers is strongly related to the GLMM definition. We have considered a GLMM (see Pinheiro & Bates, 2000 and Goldstein, 2003) to model binary response of grouped data.

Let Y_{ij} be the binary outcome of subject i of j th group, and p_{ij} the related probability of success. A GLMM could be written in the following way:

$$\text{logit}(\mathbb{E}[Y_{ij}|b_j]) = \text{logit}(p_{ij}) = \beta_0 + \sum_k \beta_k x_{ijk} + b_j,$$

where x_{ijk} are significant covariates; $b_j \sim \mathcal{N}(0, \sigma_b^2)$ are additive random effects normally distributed. The first two terms of the linear predictor ($\beta_0 + \sum_k \beta_k x_{ijk}$) are commonly called fixed effect. We fitted a GLMM on in-hospital survival outcome; the admitting hospital is the grouping factor assumed as an additive random term with normal distribution. Estimates for fixed effects coefficients (β) and

standard deviation of normal random effect (σ_b^2) can then be obtained through maximization of likelihood function

$$L(\beta, \sigma_b) = \prod_j \int \prod_i f(y_{ij}|\beta, \sigma_b, b_j)\pi(b_j)db_j,$$

where $\pi(b_j)$ is the normal density function. This integral does not have a closed form except for normal outcomes, then approximations have to be computed. We fitted GLMM models on our data with lme4 package, which makes use of Laplace approximation for computing high-dimensional integrals (Bates & Maechler, 2010).

Once we obtain the estimates of fixed effects $\hat{\beta}$ and random effect variance $\hat{\sigma}_b^2$, agglomerative algorithms (e.g. k -means) on the estimated values of random effect for each hospital can be implemented to detect clustering structure, i.e. to establish how many groups, in terms of suitable similarity indexes, might exist.

5.3 NPML estimator for GLMM

In modelling overdispersed and grouped data, the use of a fully parametric model for random effects could be too restrictive, so we have considered also the idea of NPML estimation (see Aitkin, 1999) for distribution of random effect. This idea is based on replacing normal random effect by a finite sum of mass points z_k with masses π_k . The study of z_k and of the related probabilities indicating that the amount of outcomes issued by component k is an alternative procedure to classify different hospitals. The algorithm is based on replacing integrals over normal b_j by a finite sums over K Gaussian quadrature mass point z_k with masses π_k . Then estimates of masses and fixed effects coefficients can be obtained through the maximization of

$$L(\beta, \sigma_b) = \sum_{k=1}^K \pi_k \prod_{i=1} f_{ik},$$

where $f_{ik} = \prod_j f(y_{ij}|\beta, \sigma_b, z_k)$.

The likelihood is thus a finite mixture of exponential family densities with mixture proportions π_k at mass points z_k , with the linear predictor for the ij th observation in the k th mixture component as follows:

$$\eta_{ijk} = \beta' x_{ij} + z_k.$$

The score equations turn out to be a weighted version of the single distribution score equations with weights $w_{ik} = \pi_k f_{ik} / \sum_l \pi_l f_{il}$. The estimates of these weights can be interpreted as posterior probabilities that the observation y_i comes from component k . We computed NPML estimations on our data with npmlreg package (Einbeck et al., 2009)

6. Statistical analysis

We will focus our readings on patients who have undergone primary angioplasty. We counted out patients having ‘transferred’ as a way of admittance because, concerning treatment times, they represent a different population compare with all other patients. So the population considered for the following analyses consists in 536 statistical units (patients) treated in 17 hospitals.

We have fitted a GLM and a GLMM on survival outcome. In the latter case, the admitting hospital is the grouping factor assumed as an additive random term with normal distribution.

In order to choose significant covariates for the model, we considered stepwise regression methods (Akaike information criterion) on the fixed effect part of the model and clinical best practice. These criteria pointed out Killip (P value = 0.0038) and age (P value = 8.64×10^{-5}) of patient as significant factors in order to explain survival probability, whereas clinical best practice and operational research motivation made us include among predictors Symptom Onset to Balloon time in the log scale ($\log(\text{OB})$) (P value = 0.1814) (see also Rathore *et al.*, 2009). The killip variable is now a binary categorization of Killip class, whose values are zero for less severe (Killip class 1 or 2) and one for more severe (Killip class 3 or 4) infarction. Moreover, age and $\log(\text{OB})$ have been centred.

The choice of these covariates is confirmed also in a Bayesian framework as it is explained in details in Guglielmi *et al.* (2010).

Therefore, calling Y_{ij} the binary random variable representing in-hospital survival of patient $i = 1, \dots, 536$ treated in hospital $j = 1, \dots, 17$, the models are, respectively:

$$\text{GLM} : \text{logit}(\mathbb{E}[Y_{ij}]) = \text{logit}(p_{ij}) = \beta_0 + \beta_1 \text{age}_i + \beta_2 \log(\text{OB})_i + \beta_3 \text{Killip}_i \quad (6.1)$$

$$\text{GLMM} : \text{logit}(\mathbb{E}[Y_{ij}|b_j]) = \text{logit}(p_{ij}) = \beta_0 + \beta_1 \text{age}_i + \beta_2 \log(\text{OB})_i + \beta_3 \text{Killip}_i + b_j, \quad (6.2)$$

where, in (6.2), $b_j \sim \mathcal{N}(0, \sigma_b^2)$ is the normal random effect of the grouping factor (i.e. hospital where i th patient is hospitalized). For details about this model see Ieva & Paganoni (2010).

6.1 Classification analysis

Once a model on MOMI² data has been fitted to explain in-hospital survival by means of suitable patients' covariates and process indicators, it is time to point out comparative analyses on hospitals performances.

We will now propose and compare three different approaches to catch some clustering structure in the hospital performances:

- (1) Let us compute the SSR_j for every hospital; the expected survival probability \hat{p}_{ij} is estimated with a GLM for survival outcomes with centred age, centred $\log(\text{OB})$ and Killip according to the model in (6.1). This procedure enables us to separate the hospitals in two groups: we call 'first type' the institutions whose SSR is ≥ 1 , and 'second type' the remaining ones. Only hospitals 2, 3, 5, 9, 10, 13 are clustered in the second type category, and the 12 remaining medical institutions are in the first type category.

We compute confidence intervals of nominal level 95% for the expected value of $D_j = \sum_{i=1}^{n_j} y_{ij}^{\text{obs}}$ in each hospital. In particular, the lower and upper bounds of the interval are

$$[D_j \{1 - (9D_j)^{-1} - (1.96/3)D_j^{-0.5}\}^3; D_j \{1 - (9(D_j + 1))^{-1} - (1.96/3)(D_j + 1)^{-0.5}\}^3].$$

In our case, all the confidence intervals contain the corresponding \hat{E}_j (see Fig. 2), showing the absence of outliers. For this reason, the estimation of possible clustering structure seems to suit better our data.

- (2) Fitting the model (6.2), we obtain estimates of additive contribution of each hospital to estimated survival probability. Let us call them \hat{b}_j , $j = 1, \dots, 17$.

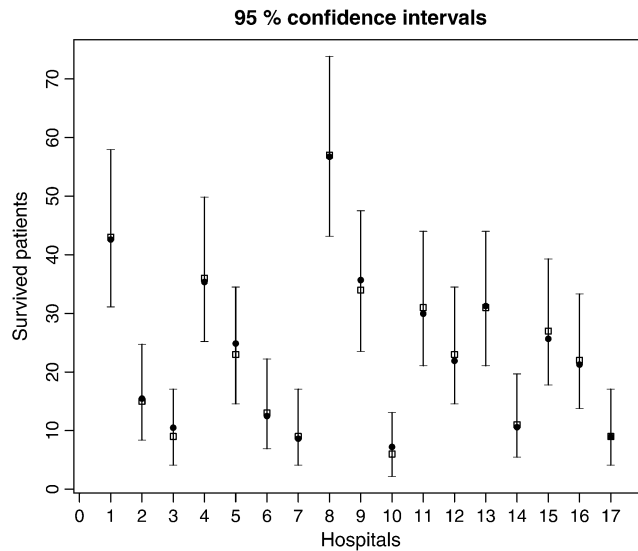


FIG. 2. Total observed number of survived patients (square), sum of survival probabilities estimated by GLM (bullet) and 95% confidence intervals for the expected value of total number of survived patients in each hospital.

TABLE 1 *Model parameters estimates and relatives asymptotic confidence intervals*

		Estimate asymptotic CI (95%)	
Intercept	$\hat{\beta}_0$	5.8551	[4.149,7.561]
Age	$\hat{\beta}_1$	-0.105	[-0.157,-0.052]
log(OB)	$\hat{\beta}_2$	-0.402	[-0.986,0.182]
Killip	$\hat{\beta}_3$	-1.719	[-2.885,-0.553]
Standard deviation	$\hat{\sigma}_b$	0.261	/

In Table 1, estimates of fixed effects coefficients and standard deviation of random effect are reported with corresponding 95% confidence intervals.

We then apply a k -means clustering algorithm (see [Hartigan & Wong 1979](#)) to the set of the \hat{b}_j s. A robustness analysis for the number of clusters using the average silhouette width ([Struyf et al., 1996](#)) supported the optimum choice of $k = 2$. Indeed, Fig. 3 shows the silhouette plot of this clustering procedure, and the value of average silhouette width equal to 0.67 indicates that a reasonable clustering structure has been found.

The means of the two clusters are 0.03456 and -0.0875 , representing a clustering structure similar to the one obtained using the previous methodology. Here again, then, we will label the two groups as first type and second type, respectively. Then we will analyse the two groups detected by this agglomerative clustering algorithm. In fact, in this case, only hospitals 2, 3, 5, 9, 10 belong to cluster with center equal to -0.0875 and for this reason, the result of this clustering procedure agrees with the previous one, except for hospital 13.

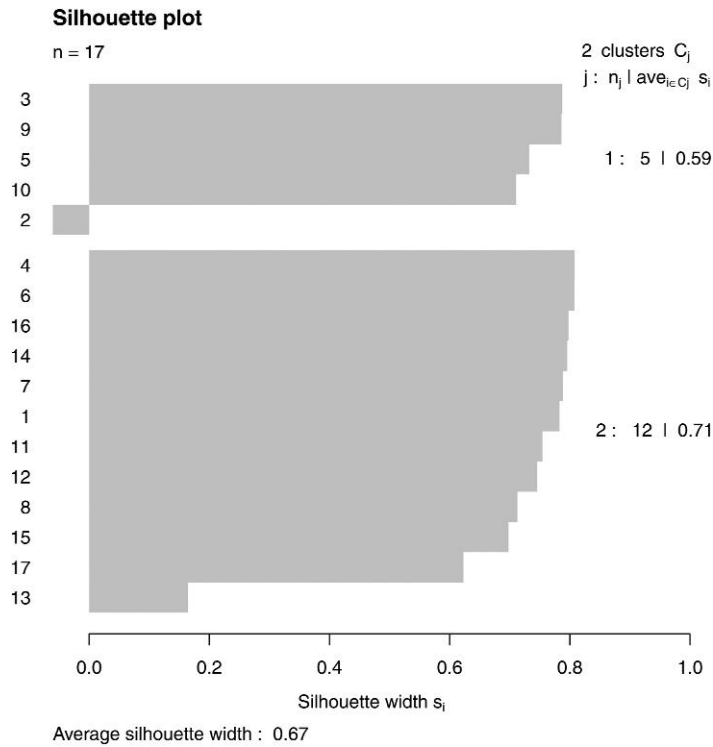


FIG. 3. Silhouette plot.

It is possible to construct a score, based on the point estimates of the hospitals' random effect. Let

$$\tilde{b}_j = \frac{\hat{\beta}_0 + \hat{b}_j}{\bar{\beta}_0},$$

where $\bar{\beta}_0$ is the empirical mean of $\hat{\beta}_0 + \hat{b}_j$. Then \tilde{b}_j is the estimate of logit of survival probability for a patient with mean age, mean symptom onset to balloon time and less severe infarction treated in hospital j , divided by the empirical mean of these survival probabilities in all the 17 hospitals. This score represents the deviation of each hospital from the overall mean of hospitals' effect. In particular, 5, 9, 3, 10, 2, 13, 17, 8, 1, 7, 14, 6, 4, 16, 11, 12, 15 is the increasing order of hospital effect, obtained by ordering the \tilde{b}_j s.

- (3) We will fit now a non-parametric GLMM, with the selected covariates of the GLMM model and with two mass points, according to the previous optimum choice ($k = 2$). Starting from this step, we classify again an hospital as first type or second type according to the arg-max of the posterior probabilities of each structure estimated for the two masses (i.e. assigning each hospital to the group whose estimated posterior probability is greater). In this case, hospital 2, 3, 5, 9, 10, 13 are clustered in the lower level, and this procedure is in total agreement with the one described in point (1). Figures 4 and 5 show the estimated survival probabilities surfaces corresponding to the two mass points (left and right panels) for increasing values of centred age and log(OB) in

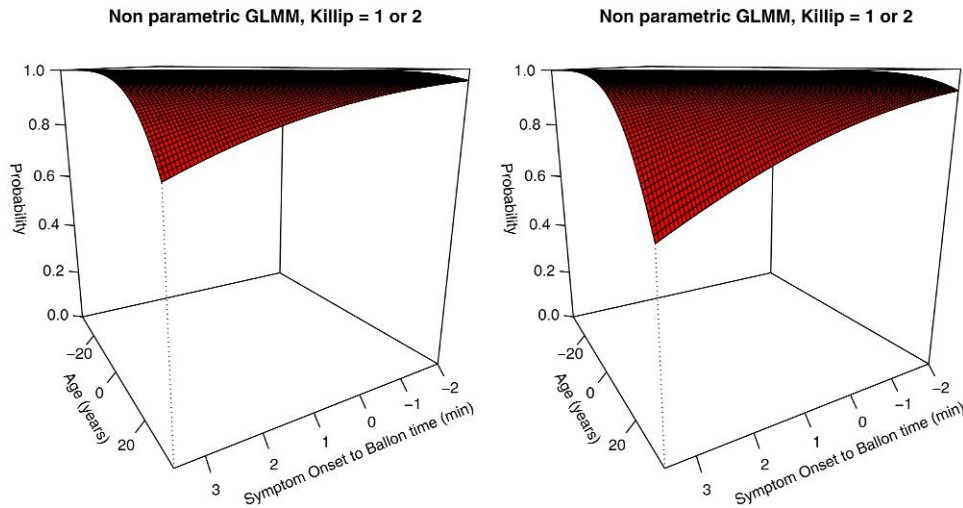


FIG. 4. Left panel: estimated survival probability surface for less severe class of Killip: *First type* hospital. Right panel: estimated survival probability surface for less severe class of Killip: *Second type* hospital.

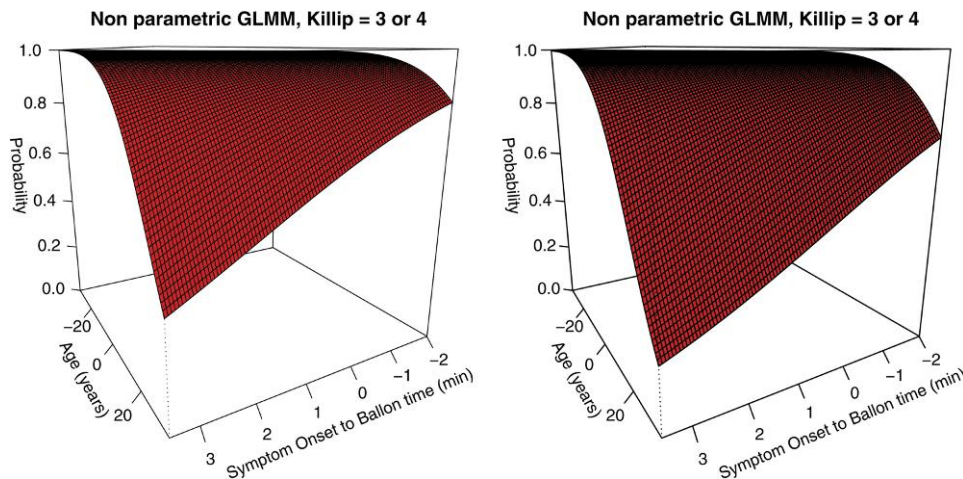


FIG. 5. Left panel: estimated survival probability surface for more severe class of Killip: *First type* hospital. Right panel: Estimated survival probability surface for more severe class of Killip: *Second type* hospital.

case of less severe (Fig. 4) and more severe (Fig. 5) infarction, respectively. These results enable the statistician to compare the different consequences of belonging to the first type or the second type on survival outcome. Here again case-mix has been considered.

In conclusion, following these three clustering procedures, we obtain the same clustering structure except for one structure (hospital 13) which is classified as second type, following the procedure (1) and (3) and as first type following the procedure (2). We can summarize results obtained with different classification techniques by comparing their performances as it is shown in Table 2.

TABLE 2 *Hospital classification using SSR index, GLMM and NPML*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
SSR	+	-	-	+	-	+	+	+	-	-	+	+	-	+	+	+	+
GLMM	+	-	-	+	-	+	+	+	-	-	+	+	+	+	+	+	+
NPML	+	-	-	+	-	+	+	+	-	-	+	+	-	+	+	+	+

The nearly unanimous agreement in the classification of the three methods support the idea that a real classification structure in two groups exists.

7. Conclusions and open problems

In this paper, we have shown how different modelling techniques can be employed in provider profiling. The results of this study support the idea of using performance indicators for comparing institutional offers of care. Measuring performance indicators is straightforward, but their relationship to actual health outcomes is often difficult to quantify. For this reason, the role of the statistician in explaining outcomes by means of suitable predictors and performance indicators becomes crucial. We have shown relatively simple and effective methods to gain these goals, and we believe that this approach could be taken into account by people in charge of health-care governance in order to support decisions in clinical context.

In general, the choice of a classifying method is effective once a model for describing the data has been chosen. In our case, GLMM and NPML methods of estimation are more suitable because they enable the statistician to take into account overdispersion of outcome induced by grouped nature of data. For this reason GLMM and NPML classification criteria should be preferred. And yet SSR criterion is easier to compute and its interpretation results to be more manageable, especially when interactions with audience coming from different background and knowhow are requested.

In this work, we have used data coming from MOMI² clinical survey on STEMI and we have modelled survival outcome by means of suitable patient's covariates and process indicators; we have proposed a ranking of hospitals and different procedures to cluster clinical institutions in groups of 'similar behaviour'. In so doing, information coming from clinical registries is used to make the hospital network more effective, improving the overall health-care process, as it is required by health-care decision makers of Regione Lombardia.

Finally, future developments of this work are comparisons of our results with those from a Bayesian classification setting (Guglielmi *et al.* 2010) and validation of these results on more complex and wider databases, coming from integration of clinical registries and administrative databases, like STEMI Archive and Public Health Database of Lombardia Region (Barbieri *et al.* 2010).

REFERENCES

- AITKIN, M. (1999) A general maximum likelihood analysis of variance components in generalized linear models. *Biometrics*, **55**, 117–128.
- ANTMAN, E. M., HAND, M., AMSTRONG, P. W., BATES, E. R. & GREEN, L. A. (2008) Update of the ACC/AHA 2004 guidelines for the management of patients with ST elevation myocardial infarction. *Circulation*, **117**, 269–329.
- BARBIERI, P., GRIECO, N., IEVA, F., PAGANONI, A.M. & SECCHI, P. (2010) Exploitation, integration and statistical analysis of Public Health Database and STEMI archive in Lombardia Region. *Complex Data Modeling and Computationally Intensive Statistical Methods*. Springer: Contribution to Statistics.

- BATES, D. & MAECHLER, M. (2010) lme4: linear mixed-effects models using S4 classes. Available at <http://CRAN.R-project.org/package=lme4>.
- 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. *J. Am. Med. Assoc.*, **283**, 2941–2947.
- DALBY, M., BOUZAMONDO, A., LECHAT, P. & MONTALESCOT, G. (2003) Transfer for primary angioplasty versus immediate thrombolysis in AMI: a meta-analysis. *Circulation*, **108**, 1809–1814.
- EINBECK, J., DARNELL, R. & HINDE, J. (2009) npmlreg: nonparametric maximum likelihood estimation for random effect models [online]. Available at <http://CRAN.R-project.org/package=npmlreg>.
- GOLDSTEIN, H. (2003) *Multilevel Statistical Models*. London: Arnolds.
- GRIECO, N., CORRADA, E., SESANA, G., FONTANA, G., LOMBARDI, F., IEVA, F., PAGANONI, A. M. & MARZEGALLI, M. (2008) Predictors of reduction of treatment time for ST-segment elevation myocardial infarction in a complex urban reality. The MOMI² survey. *Mox Report no. 10/2008*, Dipartimento di Matematica, Politecnico di Milano, Milano. Available at <http://mox.polimi.it/progetti/pubblicazioni/quaderni/10-2008.pdf>.
- GUGLIELMI, A., IEVA, F., PAGANONI, A. M. & RUGGERI, F. (2010) A Bayesian random-effects model for survival probabilities after acute myocardial infarction. *Mox Report no. 17/2010*. Dipartimento di Matematica, Politecnico di Milano, Milano. Available at <http://mox.polimi.it/progetti/pubblicazioni/quaderni/17-2010.pdf>.
- HARTIGAN, J. A. & WONG, M. A. (1979) A *k*-means clustering algorithm. *Appl. Stat.*, **28**, 100–108.
- HASDAY, D., BEHAR, S. & WALLENTIN, L. (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 (euro heart survey ACS). *Eur. Heart J.*, **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/progetti/pubblicazioni/tesi/ieva.pdf>.
- IEVA, F. & PAGANONI, A. M. (2009) A case study on treatment times in patients with ST-segment elevation myocardial infarction. *Mox Report no. 05/2009*. Dipartimento di Matematica, Politecnico di Milano, Milano. Available at <http://mox.polimi.it/progetti/pubblicazioni/quaderni/05-2009.pdf>.
- IEVA, F. & PAGANONI, A. M. (2010) Multilevel models for clinical registers concerning STEMI patients in a complex urban reality: a statistical analysis of MOMI² survey. *Commun. Appl. Indust. Math.*, **1**, 128–147.
- LINDSAY, B. G. (1983) The geometry of mixture likelihoods: a general theory. *Ann. Stat.*, **11**, 86–94.
- MCCULLAGH, P. & NELDER, J. A. (2000) *Generalized Linear Models*. London: Chapman & Hall/CRC.
- OLTRONA, L., MAFRICI, A., MARZEGALLI, M., FIORENTINI, C., PIROLA, R. & VINCENTI, A. (2005) La gestione della fase iperacuta dell'infarto miocardico con soprasslivellamento del tratto ST nella Regione Lombardia (GestIMA). *Italian Heart J.*, **6**, 489–497.
- PINHEIRO, C. & BATES, D. M. (2000) *Mixed-Effects Models in S and S-Plus*. New York: Springer.
- R DEVELOPMENT CORE TEAM (2009) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. 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. *J. Am. Stat. Assoc.*, **105**, 48–58.
- RATHORE, S. S., CURTIS, J. P. & CHEN, U. J. (2009) Association of door to balloon time and mortality in patients admitted to hospital with ST-elevation myocardial infarction: national cohort study. *Br. Med. J.*, **338**, b1807.
- SAIA, F., MARZOCCHI, A., MANARI, G., GUASTARROBA, P., VIGNALI, L. & VARANI, E. (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.

- STRUYF, A., HUBERT, M. & ROUSSEEUW, P. J. (1996) Clustering in an object-oriented environment. *J. Stat. Softw.*, **1**.
- TING, H. H., KRUMHOLTZ, H. M., BRADLEY, E. H., CONE, D. C. & CURTIS, J. P. (2008) Implementation and integration of prehospital ECGs into system of care for acute coronary syndrome. *Circulation*, **118**, 1066–1079.