MOX–Report No. 26/2013

Detecting and visualizing outliers in provider profiling via funnel plots and mixed effect models

Ieva, F.; Paganoni, A.M.

MOX, Dipartimento di Matematica "F. Brioschi"
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

mox@mate.polimi.it http://mox.polimi.it

Detecting and visualizing outliers in provider profiling via funnel plots and mixed effect models

Francesca Ieva[♯] and Anna Maria Paganoni[♯]

May 26, 2013

[♯] MOX– Modellistica e Calcolo Scientifico
Dipartimento di Matematica "F. Brioschi"
Politecnico di Milano
via Bonardi 9, 20133 Milano, Italy

`francesca.ieva@polimi.it, anna.paganoni@polimi.it`

Keywords: Funnel Plots, Generalized Linear Mixed Models, In-hospital survival, Provider Profiling, Acute Myocardial Infarction

Abstract

In this work we propose the use of a graphical diagnostic tool (the funnel plot) to detect outliers among hospitals that treat patients affected by Acute Myocardial Infarction (AMI). We consider an application to data on AMI hospitalizations arising from administrative databases. The outcome of interest is the in-hospital mortality, a variable indicating if the patient has been discharged dead or alive. We then compare the results obtained by graphical diagnostic tools with those arising from fitting parametric mixed effects models to the same data.

1 Introduction

It has become common to adopt a hierarchical model structure when comparing the performance of healthcare providers. It is not immediately clear, however, how unusual providers, that is, any with particularly high or low rates, can be identified based on such a model. Outlier detection in provider profiling is a growing interest topic in decisional processes related to healthcare regulation [2, 3, 9, 13, 15]. In fact, measuring and understanding process variability and its impact on principal outcomes are necessary for a significant improvement of healthcare systems. We focus our readings on the patterns of care undertaken by patients affected by Acute Myocardial Infarction (AMI) and hospitalized in a structure of our regional district (Regione Lombardia) in Italy. The nature of such data is hierarchical (patients within hospitals). The outcome of interest is the in-hospital mortality, a binary random variable that indicates if the patient has been discharged alive or not. We fit

suitable hierarchical logistic parametric models to these data with a twofold aim: (a) to estimate the hospitals Mortality Rates and Standardized Mortality Ratios and (b) to compute estimates of the random effects, after adjusting for significant covariates. The joint use of funnel plots and diagnostic graphical tools on the random effects estimates, leads to an effective way for detecting outliers in terms of hospital performance. These graphical tools provide an effective way to summarize information to be shared and communicate to clinicians. Moreover, they enhance possibility to convey key conclusions for supporting decision makers in the healthcare regulation process. In particular we mainly refer to [17], [11] and [18] for the use of funnel plots in healthcare regulation and to [8] for the use of graphical methods and diagnostic analysis on random effect estimates.
The article is structured as follows: first we present the data we are dealing with (Section 2), then we present suitable funnel plots to detect outliers and extreme structures with respect to Standardized Mortality Ratio (Section 3). The fitted mixed effects models and the study of random effects estimates is detailed in Section 4. In Section 5 we discuss the achieved results. All the analyses have been performed with `R` program [14].

2 Data and Extraction Criteria

Administrative health care databases play today a central role in the evaluation of healthcare systems, because of their widespread diffusion and low cost information they provide. There is an increasing agreement among epidemiologists on the validity of disease and intervention registries based on administrative databases (see, for example, [4, 5] and [19] and references therein). So more and more frequently administrative data are used to address epidemiological issues in observational studies. The most critical issue when using administrative databases for observational studies is represented by the selection criteria of the statistical units. In fact several different criteria may be used, and they will result in different images of prevalence or incidence of diseases. In the case of interest, we focus on in-hospital mortality after an Acute Myocardial Infarction (AMI). Concerning this pathology, since every hospital admission ends in a record collected in the administrative datawarehouse, the administrative database of SDO (*Scheda di Dimissione Ospedaliera*, i.e., hospital discharge paper) has been used in order to identify AMI episodes and related subsequent hospitalizations. In fact, the SDO database contains data for each hospitalization that a patient experiences along time, providing information both on patient features (in terms of sex, age, ...) and on her/his hospitalization details (date of admission and discharge, diagnoses and procedures, type of admission, type of discharge, vital status at discharge, hospital of admission/discharge, ...).
The case study presented here concerns data arising from a project named "Exploitation, integration and study of current and future health databases in Lombardia for Acute Myocardial Infarction", funded by Ministry of Health and Regione Lombardia, the region in the northern part of Italy whose capital is Milan. The principal

aim of the project was to exploit clinical registries and administrative databases for evaluating the performance of the cardiological network in treating AMI patients and to carry out an effective healthcare planning based on real evidence and needs. In order to identify AMI patients within the administrative database we considered admissions ended between 2000 and 2010 and classified within the Major Diagnostic Categories (MDC) 01 - *Nervous System*, 04 - *Respiratory System* and 05 - *Circulatory System*. Among these records, admissions for acute myocardial infarction have been identified as those presenting an AMI code in the first diagnosis field among the six available in the SDO. The list of ICD-9-CM codes referred to AMI has been created following AHRQ-IQI [1] and consists in the following ones: 41000; 41001; 41010; 41011; 41020; 41021; 41030; 41031; 41040; 41041; 41050; 41051; 41060; 41061; 41070; 41071; 41080; 41081; 41090; 41091. Patients under 18 years, and patients transferred by another hospital have not been considered. We do not consider also hospitals managing a number of cases less than 20.
The dataset consists then of 46079 patients treated in 96 hospitals of Regione Lombardia.

3 Funnel plots for Standardized Mortality Ratio

The first statistical method we consider for outliers detection is based on funnel plots. The funnel plots have been originally introduced in meta-analysis studies, with the primary aim of being a visual aid to detection of bias or systematic heterogeneity. More recently they have been proposed as a graphical aid for institutional comparisons, especially because they overcome some criticism of the more traditional caterpillar plots [12].
Following [11] we consider three different approaches to identify unusual performance of providers using the administrative data previously introduced: (a) a funnel plot based on the common mean model, (b) a funnel plot to identify outliers in the random effects distribution and (c) a funnel plot to identify extremes in the random effects distribution. In the construction of a funnel plot we need, an indicator of performance X, a target θ_0 which represents the desired expectation, so that $\mathbf{E}[X] = \theta_0$ for institution "in-control" and a precision parameter ρ that controls the accuracy of the measured indicator. In general ρ is inversally proportional to the variance, i.e., $\rho \propto \frac{1}{Var[X]}$. Given a series of J observations x_j, and the associated precisions ρ_js a funnel plot is a scatterplot of x_j versus ρ_j, with superimposed control limit lines, function of ρ, computed according to the chosen model for accounting overdisperion. In our case we chose as performance indicator of each hospital the Standardized Mortality Ratio (SSR) defined as

$$SSR_j = \frac{\sum_{i=1}^{n_j} y_{ij}^{obs}}{\sum_{i=1}^{n_j} \hat{p}_{ij}} = \frac{O_j}{E_j},\ j = 1, ..., J. \tag{1}$$

where y_{ij}^{obs} is the observed outcome of in-hospital mortality for patient i treated in the hospital j, n_j is the number of patients treated in hospital j and $\hat{p}_{ij}$ is the

corresponding estimated probability of death.
SSR_j relates the actual mortality at the jth hospital to the expected mortality in the same hospital, adjusted for different patient severity resumed in the covariates of the logistic model. The estimated number of deaths E_j, can be obtained fitting the following logistic regression model:

$$logit\,(\mathbb{E}[Y_i]) = logit(p_i) = \beta_0 + \sum_{k=1}^{2} \beta_k x_{ik}. \quad (2)$$

Y_i is the Bernoulli variable representing the in-hospital mortality of patient i and $(\beta_0, \beta_1, \beta_2)$ are the parameters corresponding to the fixed covariates: $\mathbf{x}_i = (x_{1i}, x_{2i}) =$ (Age$_i$, Risk Index$_i$). We considered also the sex of the patient, but it has been discarded because highly correlated with the age. In the selected model both the covariates have a very high statistical significance (p-value $< 2 * 10^{-16}$) and the Risk index evaluated with the APR-DRG mortality score see [6, 16]. Finally, as a precision parameter to be displayed on the abscissa we adopted the expected number of deaths of each hospital.
Starting from the first case (a) which doesn't consider any overdispersion due to the grouped nature of data, we draw a funnel plot for Standardized Mortality Ratio reported in Figure 1. In this case limits are given by

$$\theta_0 \pm z_p \sqrt{\frac{\theta_0}{E}} \quad (3)$$

where θ_0 is the fixed target (in our case θ_0 is set to 1 meaning that the target level is a number of observed deaths which is equal to the expected one), z_p are the quantiles of a standard gaussian distribution, and E is the volume of expected cases.
In case (b) an additive adjustment factor is introduced in the band limits of the funnel plot (see Figure 2). The limits are then given by

$$\theta_0 \pm z_p \sqrt{\frac{\theta_0}{E} + \hat{\tau^2}} \quad (4)$$

where, following [7], if we call $\sigma_j^2 = \frac{\theta_0}{E_j}$, $w_j = 1/\sigma_j^2$ for every hospital $j = 1, ..., J$, $\hat{\tau^2}$ is given by

$$\hat{\tau^2} = \frac{J\hat{\phi}^W - (J-1)}{\sum_j w_j - \sum_j w_j^2 / \sum_j w_j}$$

and $\hat{\phi}^W$ is estimated using the q-Winsorised Z-scores z_j^W

$$\hat{\phi}^W = \frac{1}{J} \sum_{j=1}^{J} (z_j^W)^2.$$

The Z-scores z_j are obtained by standarizing the observed outcomes of interest, choosing the target θ_0 as the mean, and the estimated variability $\sqrt{\sigma_j^2}$ as standard

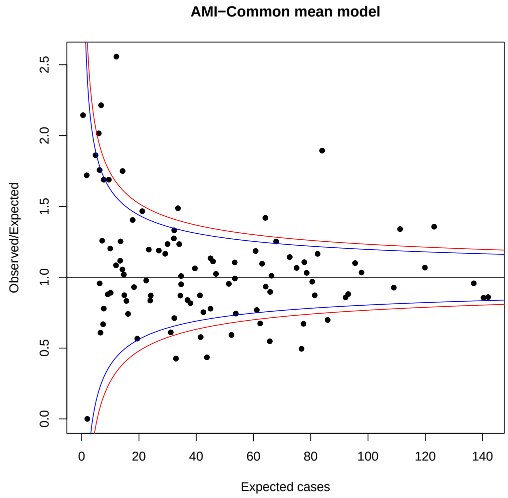

Figure 1: In-hospital Standardized Mortality Ratio from 96 hospitals in Regione Lombardia. Unadjusted funnel plot with band limits at 95% (blue lines) and 98% (red lines). The horizontal solid black line is the target limit.

deviation, i.e.,

$$z_j = \frac{SSR_j - \theta_0}{\sqrt{\mathrm{Var}[X_j]}} = \frac{SSR_j - \theta_0}{\sqrt{\frac{\theta_0}{E_j}}}.$$

The q-Winsorised Z-scores are obtained setting the lowest $100q$ Z-scores to the quantile of order q of $z_1, ..., z_J$ and, analogously the highest $100q$ Z-scores to the quantile of order $(1-q)$ of $z_1, ..., z_J$. If there is no true overdispersion $J\hat{\phi}^W$ has approximately a χ^2 distribution with J degrees of freedom. In this case $\mathbf{E}[\hat{\phi}^W] = 1$ and $\mathrm{Var}[\hat{\phi}^W] = \frac{2}{J}$. So data require a statistically significant adjustment for overdispersion when $\hat{\phi}^W > 1 + 2\sqrt{2/J}$. In our application $\hat{\phi}^W = 2.6$ and this value is strictly greater than $1 + 2\sqrt{2/J} = 1.28$. This fact proves, as suggested in [17], a statistical evidence for overdispersion.
In case (c) a multiplicative adjustment factor is introduced in the band limits of the funnel plot (see Figure 3). The limits are then given by

$$\theta_0 \pm z_p \sqrt{\frac{1}{\hat{\tau^2}} \frac{\theta_0}{E} \left(\frac{\theta_0}{E} + \hat{\tau^2} \right)}. \tag{5}$$

The plot is reported in Figure 3. As stressed in [11] the case (b) is the most appropriate adjustment for the identification of outliers to random effect distribution, and so to detect outlier institutions with unusual performance. Moreover the multiplicative adjustment proposed in (c) will tend to identify too many providers as unusual.

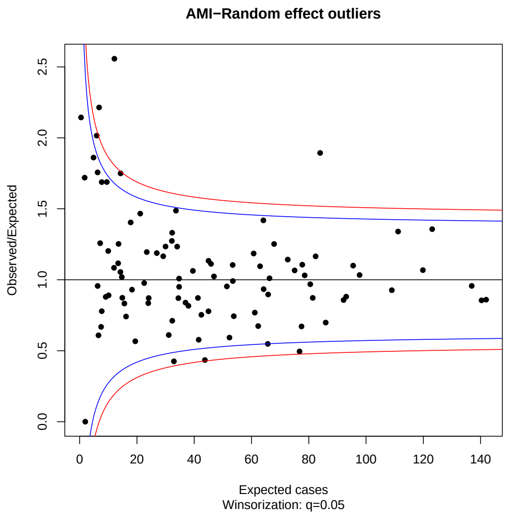

Figure 2: In-hospital Standardized Mortality Ratio from 96 hospitals in Regione Lombardia. Additive adjusted funnel plot with band limits at 95% (blue lines) and 98% (red lines). The horizontal solid black line is the target limit.

4 Outlier detection in random intercept estimates

The second statistical method we consider for outliers detection is based on the identification of extremes values of the random effect point estimates arising a suitable mixed effect model fitted to the data, where the hospital of admission is considered as a grouping factor. The study of the random effects distribution for provider profiling has used both in frequentist [9] and Bayesian context [10] with the main aim of clustering hospitals and detecting similar behaviors. In this paper we take advantage of the study of random effect distribution to understand the agreement with the graphical procedures presented in Section 3 and then to provide further evidence to decision makers, together with a proper quantification of the outlyingness of each performance.

So we fit a parametric mixed effect model to explain the in-hospital mortality, with in the fixed part the significant covariates selected in model (2) and a random intercept. In particular

$$logit\left(\mathbb{E}\left[Y_{ij}|b_j\right]\right) = logit(p_{ij}) = \\ \beta_0 + \sum_{k=1}^{2} \beta_k x_{ijk} + b_j. \tag{6}$$

Y_{ij} is the Bernoulli variable representing the in-hospital mortality of patient i treated in hospital j. $(\beta_0, \beta_1, \beta_2)$ is the parameters vector corresponding to the fixed part. $b_j \sim \mathcal{N}(0, \sigma_b^2)$ is the Normal random effect superimposed to the grouping factor (i.e., the hospital where a patient is admitted to).

In our application the point estimates of parameters ($\pm$ standard deviation) are

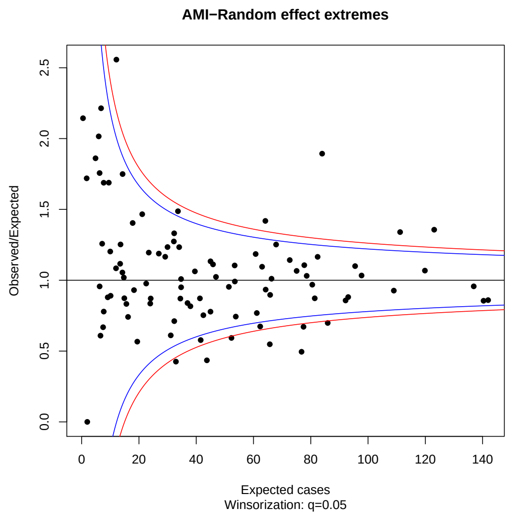

Figure 3: In-hospital Standardized Mortality Ratio from 96 hospitals in Regione Lombardia. Multiplicative adjusted funnel plot with band limits at 95% (blue lines) and 98% (red lines). The horizontal solid black line is the target limit.

$\hat{\beta}_0 = -7.8108 \pm 0.1379$, $\hat{\beta}_1 = 0.0408 \pm 0.0017$ and $\hat{\beta}_2 = 1.1257 \pm 0.0214$, respectively.

The estimate of the standard deviation of the random effect is equal to $\hat{\sigma_b} = 0.37$. The points over the limits of the whiskers in the boxplot (see Figure 4) of the point estimates of the random intercept in model (6) are a subset of the hospitals which locate out of the 95% limits of all three funnel plots in (3), (4), (5). In Figure 4 we draw also the quantiles of order 2.5 % and 97.5% (blue lines) and the quantiles of order 1% and 99% (red lines), in order to highlight the agreement with Figure 2. In Figure 5 the Normal Q-Q plot of the point estimates of the random intercept in model (6) is shown. In fact the Normality assumption on the random effect b_j in model (6) is questioned and weakened by the detected outliers in Figure 4). Moreover the detected outlier hospitals are the ones located out of the limits in the funnel plot shown in Figure 2 that, as said before is a suggested tool to point out unusual performance.

5 Conclusions

The identification of unusual performance of healthcare institutions is a hot topic of the last years. It is crucial, especially in supporting decisions of people in charge with healthcare governance. We propose the joint use of graphical instruments (like funnel plots) to detect and visualizing outliers, and the study of random effects estimated by fitting a hierarchical mixed effects model to the same data.

The use of hierarchical model allows for a more proper modelling of the overdisperion highlighted by the funnel plots. Moreover, they enable researchers to quantify the

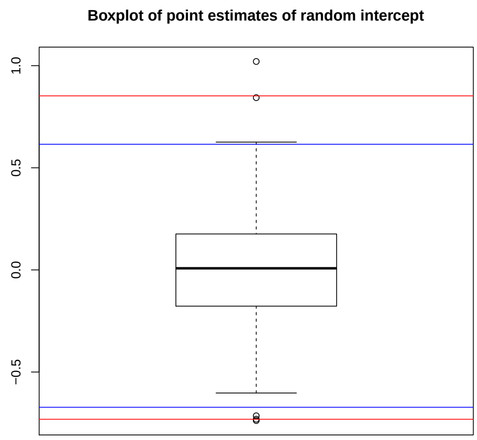

Figure 4: Boxplot of the point estimates of the random intercept in model (6). The horizontal blue lines correspond to the quantiles of order 2.5 % and 97.5%. The horizontal red lines correspond to the quantiles of order 1% and 99%.

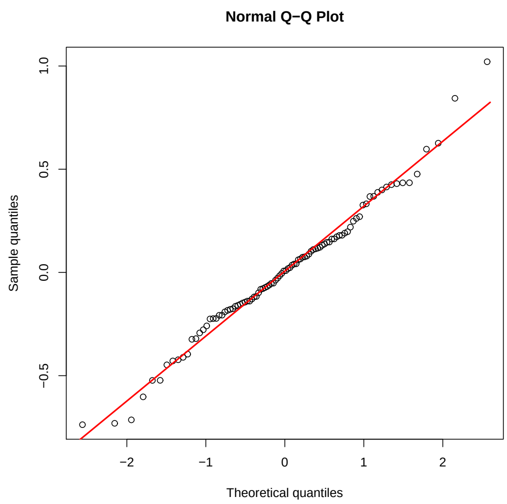

Figure 5: Normal Q-Q plot of the point estimates of the random intercept in model 6.

outlyingness of a performance, providing a quantitative support to decision makers. In fact, the funnel plots indicate the presence of overdispersion in data, and by computing suitable confidence limits they allow to detect unusual institutions. On the other hand, the study of tails and extremes in estimated random effect distribution confirm the obtained results and provide also an estimate of the quantitave effect of

these structures in outcome prediction.
We propose the joint use of these two techniques as a simple and effective way for profiling providers within the context of healthcare regulation.

Acknowledgments

This work is within the Strategic Program “Exploitation, integration and study of current and future health databases in Lombardia for Acute Myocardial Infarction”. The authors wish to thank Regione Lombardia - Healthcare division for having funded and sustained the project, Lombardia Informatica S.p.A. for having provided data and all the physicians who collaborated to STEMI Archive planning and data collection.

References

[1] AHRQ QualityIndicators (2007) Guide to Inpatient Quality Indicators: Quality of Care in Hospitals - Volume, Mortality, and Utilization, Department of Health and Human Services Agency for Healthcare Research and Quality, Version 3.1.

[2] Austin, P. C., Naylor, C. D., and Tu, J. V. (2001), A Comparison of a Bayesian vs. a Frequentist Method Profiling Hospital Performance, *Journal of Evaluation in Clinical Practice*, **7**, 3545.

[3] Austin, P. C., Alter, D. A., and Tu, J. V. (2003), The Use of Fixed- and Random-Effects Models for Classifying Hospitals as Mortality Outliers: A Monte Carlo Assessment, *Medical Decision Making*, **23**, 526539.

[4] 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 Series “Contribution to Statistics”, Springer, 41-56.

[5] Campbell, M.J., Jacques, R.M., Fotheringham, J., Maheswaran, R., Nicholl, J. (2012), Developing a summary hospital mortality index: retrospective analysis in English hospitals over five years, *British Medical Journal* **344**, e1001.

[6] De Marco, M.F., Lorenzoni, L., Addari. P., Nante. N. (2002), Evaluation of the capacity of the APR-DRG classification system to predict hospital mortality, *Epidemiologia e Prevenzione*, **26**, 4, 183-190.

[7] DerSimonian, R., Laird, N. (1986), Meta-Analysis in Clinical Trials, *Controlled Clinical Trials*, **7**, 177-188.

[8] Farrel, P.J., Groshen, S., MacGibbon, B., Tomberlin, T.J. (2010), Outlier detection for a hierarchical Bayes model in a study of hospital variation in surgical procedures, *Statistical Methods in Medical Research*, **19**, 601-619.

[9] Grieco, N., Ieva, F., Paganoni, A.M. (2012). Performance assessment using mixed effects models: a case study on coronary patient care. *IMA Journal of Management Mathematics*, **23**, 2, 117-131.

[10] Guglielmi, A., Ieva, F., Paganoni, A.M., Ruggeri, F., Soriano, J. (2013). Semiparametric Bayesian modeling for the classification of patients with high observed survival probabilities. *Journal of the Royal Statistical Society Series C.* Forthcoming.

[11] Jones, H.E., Spiegelhalter, D.J. (2011), The Identification of "Unusual" Health-Care Providers From a Hierarchical Model, *The American Statistician*, **65**, 3, 154–163.

[12] Mohammed, M.A., Deeks, J.J. (2008), In the context of performance monitoring, the caterpillar plot should be mothballed in favor of the funnel plot, *The Annals of Thoracic Surgery*, **86**, 348.

[13] Normand, S. L. T., Glickman, M. E., and Gatsonis, C. A. (1997), Statistical Methods for Profiling Providers of Medical Care: Issues and Applications, *Journal of the American Statistical Association*, **92**, 803814.

[14] R Development Core Team (2009), R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. [Online] http://www.R-project.org

[15] 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**, 489, 48–58.

[16] Romani, P.S., Chan, B.K. (2000), Risk-adjusting acute myocardial infarction mortality: are APR-DRGs the right tool? *Health Services Research*, **34**, 7, 1469–1489.

[17] Spiegelhalter, D.J. (2005), Funnel plots for comparing institutional performance, *Statistics in Medicine*, **24**, 1185-1202.

[18] Spiegelhalter, D.J., Sherlw-Johnson, C., Bardsley, M., Blunt, I., Wood, C., Grigg, O. (2012), Statistical methods for healthcare regulation: rating, screening and surveillance, *Journal of the Royal Statistical Society - Series A*, **175**, 1-47.

[19] Wirehn, A.B., Karlsson, H.M., Cartensen J.M., *et al.* (2007) Estimating Disease Prevalence using a population-based administrative healthcare database. *Scandinavian Journal of Public Health*, 35: 424–431

MOX Technical Reports, last issues

Dipartimento di Matematica "F. Brioschi",
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

25/2013 Cattaneo, Laura; Zunino, Paolo
Computational models for coupling tissue perfusion and microcirculation

24/2013 Mazzieri, I.; Stupazzini, M.; Guidotti, R.; Smerzini, C.
SPEED-SPectral Elements in Elastodynamics with Discontinuous Galerkin: a non-conforming approach for 3D multi-scale problems

26/2013 Ieva, F.; Paganoni, A.M.
Detecting and visualizing outliers in provider profiling via funnel plots and mixed effect models

23/2013 Srensen, H.; Goldsmith, J.; Sangalli, L.M.
An introduction with medical applications to functional data analysis

22/2013 Falcone, M.; Verani, M.
Recent Results in Shape Optimization and Optimal Control for PDEs

21/2013 Perotto, S.; Veneziani, A.
Coupled model and grid adaptivity in hierarchical reduction of elliptic problems

20/2013 Azzimonti, L.; Nobile, F.; Sangalli, L.M.; Secchi, P.
Mixed Finite Elements for spatial regression with PDE penalization

19/2013 Azzimonti, L.; Sangalli, L.M.; Secchi, P.; Domanin, M.; Nobile, F.
Blood flow velocity field estimation via spatial regression with PDE penalization

18/2013 Discacciati, M.; Gervasio, P.; Quarteroni, A.
Interface Control Domain Decomposition (ICDD) Methods for Coupled Diffusion and Advection-Diffusion Problems

17/2013 Chen, P.; Quarteroni, A.
Accurate and efficient evaluation of failure probability for partial different equations with random input data