MOX–Report No. 5/2009

A case study on treatment times in patients with ST-Segment Elevation Myocardial Infarction

Francesca Ieva, Anna Maria Paganoni

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

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

A case study on treatment times in patients with ST-Segment Elevation Myocardial Infarction

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

19th January 2009

[a] MOX– Modellistica e Calcolo Scientifico
Dipartimento di Matematica "F. Brioschi"
Politecnico di Milano
via Bonardi 9, 20133 Milano, Italy
francesca.ieva@fastwebnet.it
anna.paganoni@polimi.it

Keywords: Generalized Linear Models, Generalized Additive Models, Contingency Tables, Odds Ratio.

Abstract

In this paper we conduct a statistical analysis of data coming from an observational case study about patients with ST-Segment Elevation Acute Myocardial Infarction treated in one of the 23 hospitals of the Milano network for acute coronary syndromes and emergency services. The principal aim of this article is to identify from a statistical perspective the most important prognostic factors for in-hospital survival and reperfusion efficacy. We model the dependency between outcome variables and predictors with Generalized Additive Models. These statistical analyses have demonstrated the clinical guess that an early pre-alarm of the Emergency Room is an essential step to improve the clinical treatment of patients.

1 Introduction

The Working Group for Cardiac Emergency in Milano, the Cardiology Society, and the 118 Dispatch Center (national free number for medical emergencies) have considered as a mandatory step to streamline an optimal care process for patients with ST-segment Elevation Myocardial Infarction (STEMI), to collect and analyze enough sampling data, both at time to treatment and admittance modality of patients with STEMI, referred to one of the 23 hospitals of the Milano network for acute coronary syndromes and emergency services. So we describe here this observational case study about patients with STEMI, focusing our attention on different statistical analyses conducted in order to identify the most important prognostic factors and to model their influence on outcome variables. We analyse the data collected in a capillary study during five time periods lasting from 30 to 60 days in the urban area of Milano in $2006-2008$, called MOMI^2 (MOnth MOnitoring Myocardial Infarction in MIlan). The statistical analyses of this dataset have demonstrated the clinical guess that an early pre-alarm of the Emergency Room (ER) is an essential step to improve the clinical treatment of patients. Pre-hospital and in-hospital times are highlighted as fundamental factors we can act on to reduce the in-hospital mortality and to increase the rate of effective reperfusion treatments of infarcted related arteries. In particular we prove that, in order to make the Door to Balloon time lower than 90 minutes, i.e. the limit suggested by the AHA/ACC guidelines, it is fundamental to make and transmit the electrocardiogram as soon as possible. The statistical evidence of the usefulness of the first electrocardiogram execution, in terms of time to treatment reduction, prompted the management of 118 Dispatch Center to equip all the Basic Life Support Units with instruments for electrocardiogram teletrasmission. Moreover this study has been taken as a prototype for the design of a new statistical study which is a part of a strategical program of the Lombardia region, funded by the Ministry of Health, under the program "Exploitation, integration and study of current and future health databases in Lombardia for Acute Myocardial Infarction".

2 Materials and Methods

A net connecting the territory to 23 hospitals, by a centralized coordination of the emergency resources has been activated in the Milan urban area since 2001. Its primary aims are promoting the best utilization of the different reperfusion strategies, reducing transport and decisional delays connected with logistic matters and therapies, and increasing the number of patients undergoing primary Percutaneous Coronary Intervention (PCI) before 90 minutes since the arrival at ER (AHA guidelines, 2007). Difficulties in reaching these goals are primary due to the fact that Milan urban area is a complex territory with high density of population (2.9 million resident and 1 million commuters daily) and a great number of hospitals ($n = 27$). Twenty-three of them have a cardiology division and a Critical Care Unit; 18 offer a 24 hour available Cath Lab for primary PCI, 5 are completed with a Cardiac Surgery unit. In order to monitor network activity, time to treatment and clinical outcome, the data collection MOMI^2 on data related to patients admitted to hospitals belonging to the net was planned and made, during five period corresponding to five monthly/bimestral collection (respectively: MOMI^2.1 from jun 1st to 30th 2006, MOMI^2.2 nov 15th to dec 15th 2006, MOMI^2.3 jun 1st to jul 30th 2007, MOMI^2.4 nov 15th to dec 15th 2007, MOMI^2.5 jun 1st to 30th 2008). Information concerning demographic features (sex, age), clinical appearance (presenting symptoms, Killip class at admittance), way of admittance in hospital (spontaneous, with BLS (Basic Life Support), with ALS (Advanced Life Support), with or without ECG teletrasmission), symptom onset times, in–hospital times (first ECG times, DB (Door to Balloon) times), hospital organization (alert, Fast Track) and clinical outcomes (in-hospital mortality, ST-elevation reduction) have been listed and studied for patients treated with primary PCI, for a total of 437 statistical units. First of all we were asked for an explorative analysis of this health database. We conducted it to asses dependence patterns between covariates, to study from a statistical perspective the relationships suggested by the clinical know-how and to detect potential variables we can act on in order to improve the performance indicators of the hospitals. Then we studied the dependence between the principal performance indicators (in particular times to treatment) and the clinical outcomes. Other studies concerning this kind of problem have found conflicting results regarding the relationship between mortality and time to reperfusion with PCI. Some investigators have found lower mortality for shorter onset-to-balloon times for all patients or just certain subgroups such as high-risk patients (Cannon et al. 2000). Other studies found no lower mortality for shorter symptom onset to balloon time but did find lower mortality for shorter DB time (Jneid et al. 2008). Finally, some studies failed to find an association between mortality and pre-hospital and in-hospital times (McNamara et al. 2006). We detected a connection between outcomes and times (both concerning symptoms onset and in-hospital times); in particular our data pointed out the dependence between the efficacy of the operation (measured by reduction of ST-segment elevation) and time of DB and Symptoms onset time. We proposed to clinicians the use of a Generalized Linear Model and a Generalized Additive Model (Hastie and Tibshirani 1986, 1987, 1999) on MOMI^2 data in order to explain the outcomes of interest (in-hospital mortality and reperfusion efficacy) by means of the other suitable covariates of the dataset. Moreover, we want to give also inferential estimates of some parameters, such as odds ratios, quantifying the association between outcomes and covariates. Unfortunately the usual asymptotic inferential techniques, used to estimate the parameters of fitted generalized models such as logistic regression models, fail in this context characterized by small or unbalanced sample sizes. So we are forced to study properties and performances and then to apply some techniques to produce non asymptotic inference results on contingency tables constructed from small or unbalanced samples. In order to do this a comparison numerical study between different exact confidence intervals on the odds ratios, measuring associations between variables, was therefore conducted and implemented by statistical software `R.7.1`(Agresti 1992, 1999; Agresti and Min 2002; Gart 1962, 1969; Gart and Thomas 1972).

3 Explorative Biodata Mining

An explorative analysis of data was performed to study the dependence between the DB time and potential covariates we can act on in order both to improve the time gain in reperfusion therapy and to increase in-hospital survival (see Ting et al. 2008; Grieco et al. 2008). A CART analysis using Gini's impurity index splits groups satisfying or not the limit of 90 min for DB time in terms of way of hospital admittance and time of first ECG within or not 10 minutes (see also Ting et al. 2008), limit suggested by the AHA/ACC guidelines (see Figure 1). In fact the distribution of the DB time in the population of patients

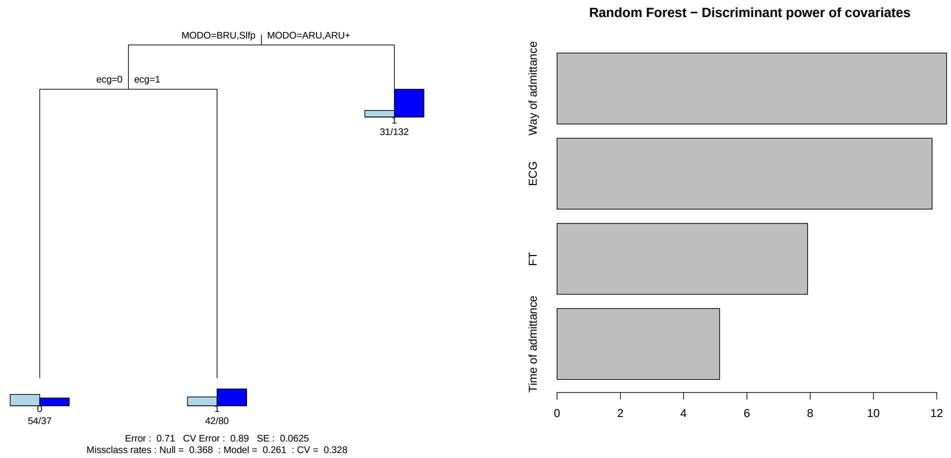

Figure 1: Left panel: CART. Right panel: Random Forest on CART predictors assessing discriminatory power of covariates.

with the first ECG within 10 minutes is confirmed to be stochastically inferior to the corresponding distribution in patients with the first ECG after 10 minutes; this stochastic order between distributions is confirmed by Mann-Whitney non parametric comparison test: p-value $< 10^{-5}$ (see Figure 2). A random forest analysis (Breiman 2001) applied to CART predictors (Breiman et al 1984) has been performed in order to asses the discriminatory power of covariates. All these signals coming from data highlight the same matter: in order to make the time of DB lower than 90 minutes, it is fundamental to make and transmit ECG as soon as possible. In fact there is a masking effect between covariates detected by the classification analysis: the way of in-hospital admittance and the time of first ECG; the exact Fisher test, performed on the contingency table of the way of hospital admittance and a variable indicating if the time of first ECG is within or not 10 minutes, shows a strongly statistical evidence (p-value $< 10^{-5}$) of dependence between these two covariates. So the main difference in ways of admittance is their possibility of sending electrocardiograms.

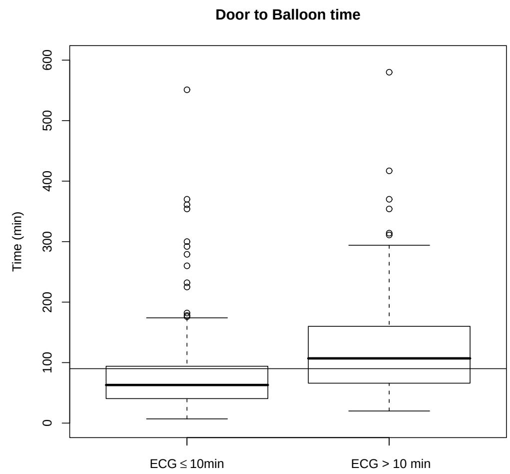

Figure 2: Flanked boxplots of distributions of the DB time in patients with the first ECG within or not 10 minutes.

4 Non parametric approach to generalized linear models

In order to explain the outcomes as function of the other variables of the dataset and to justify the effort to improve the performance indicators, we considered two different statistical techniques to model binary response: a Generalized Linear Model (GLM), and a Generalized Additive Model (GAM). The responses of interest are the in-hospital survival and the reperfusion efficiency (represented by the 70% reduction of the ST-segment elevation an hour after the PCI), which are both binary variables. First we model the probability related to the outcome variables with a linear logistic regression model; if we denote the outcome variable under study as Y, and the set of p predictors as $\mathbf{X}$, a GLM for the binary response Y can be written as

$$\mathrm{logit}\{\mathrm{P}(\mathbf{X})\} \equiv \log\left\{\frac{\mathrm{P}(\mathbf{X})}{1-\mathrm{P}(\mathbf{X})}\right\} = \alpha + \sum_{\mathrm{j}=1}^{\mathrm{p}} \beta_{\mathrm{j}} \mathrm{X}_{\mathrm{j}} \tag{1}$$

where $P(\mathbf{X}) = \mathrm{pr}(Y = 1|\mathbf{X})$. A stepwise model selection procedure based on backward selection, AIC criterion, and clinical best practice, pointed out as explanatory variables the killip class, age and total ischemic time (Symptom onset to Balloon time) for the survival outcome. On the other hand, the Door to Ballon time and the Symptom onset time has been pointed out as explanatory variables for the efficacy outcome. As an example, Figure 3 shows the result of the implemented model. The drawn surface describes the probability of an effective reperfusion; this probability decreases as long as both the Symptom onset time and the Door to Ballon time increase.

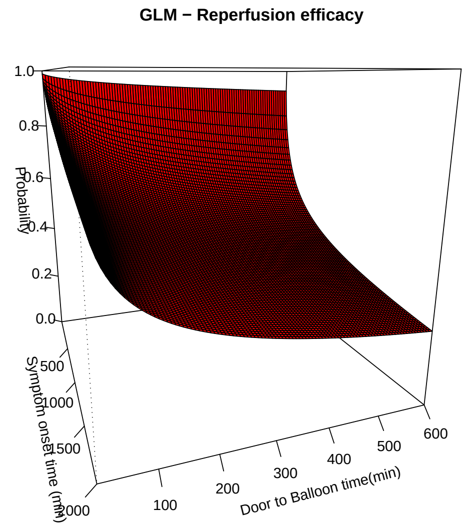

Figure 3: Reperfusion efficacy surface estimated by GLM of logistic regression.

The in-hospital survival probability decreases as long as both the age and the total ischemic time increase, for both the cases of less and more severe STEMI (measured by the killip class), but more strongly for the second one. Nevertheless, a non-parametric model would give the data more of a chance to speak for themselves without forcing the model into a rigidly defined class. So we performed the same analysis carried out with GLM, now using generalized additive models (GAM). If we denote the outcome variable under study as Y, and the set of p predictors as $\mathbf{X}$, the logistic regression in terms of generalized additive models can be written as

$$\mathrm{logit}\{\mathrm{P}(\mathbf{X})\} \equiv \log\left\{\frac{\mathrm{P}(\mathbf{X})}{1-\mathrm{P}(\mathbf{X})}\right\} = \alpha + \sum_{\mathrm{j}=1}^{\mathrm{p}} \mathrm{s_j}(\mathrm{X_j}) \qquad (2)$$

where $P(\mathbf{X}) = \mathrm{pr}(Y = 1|\mathbf{X})$ and $s_j(X_j)$ are p arbitrary unspecified univariate functions. We estimate these functions s_j with suitable smoothers and in this case we choose as smoother functions natural cubic splines of order three (see Hastie and Tibshirani 1986 for details). Even if the value of the AIC statistics for GAM fitted to explain in-hospital survival or reperfusion efficay are very similar to the values attained by the logistic regression models we can appreciate advantages of non parametric modelling in terms of flexibility and adaptation to the data. In these case we avoided to superimpose the linear link between the function of the response variable and predictors, letting data better describe their own features. As it was expected, the survival probability decreases as the predictors increase, but in the specific way imposed by data features, not only by the superimposed linear model. In Figure 4 are shown the marginal estimated smoother functions $s_j(X_j)$ of age and Symptom onset to Balloon time (OB) in the GAM model of in-hospital survival, for both the cases of less and more severe STEMI.

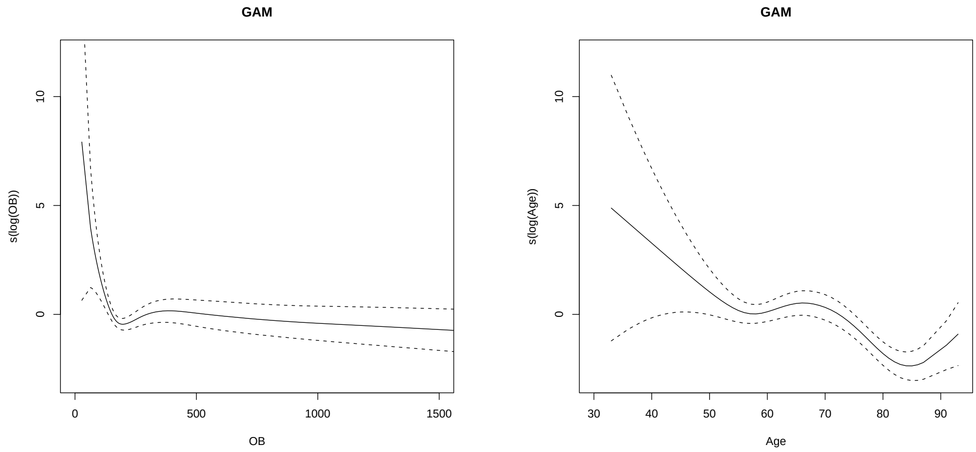

Figure 4: Estimated smoother for total ischemic time and age for GAM model of logistic regression for in-hospital survival (pointwise standard errors included).

In Figure 5 are reported the marginal estimated smoother functions $s_j(X_j)$ of Symptom onset time and Door to Balloon time in the GAM model of reperfusion efficacy. We can visualize (see Figure 6) also the estimated surface of reperfusion efficacy as function of both Symptom onset time and Door to Balloon time. In fact the robustness of the analysis have been proved by similar results obtained both with Generalized Linear Models and with Generalized Additive Models. The fact that a more sophisticated and data adaptive method has reached similar conclusions validates the results of the statistical analysis.

These results support the effort of acting on some covariates in order to attain an improvement in performance indicators (such as the reduction of DB time) and so to increase the probability of a successful treatment. Analogously it would be strongly important to persuade the population to call the free emergency number as soon as possible after the Symptom onset. In order to reach these goals, has been established by the governance of Lombardia Region to equip all the Basic Support Units covering Milan urban area with instruments for electrocardiogram teletrasmission. Moreover, awareness campaigns are planned for spring and autumn 2009 to make people sensitive about the problem and to spur them to call the free number of emergencies as soon as possible when symptoms arise.

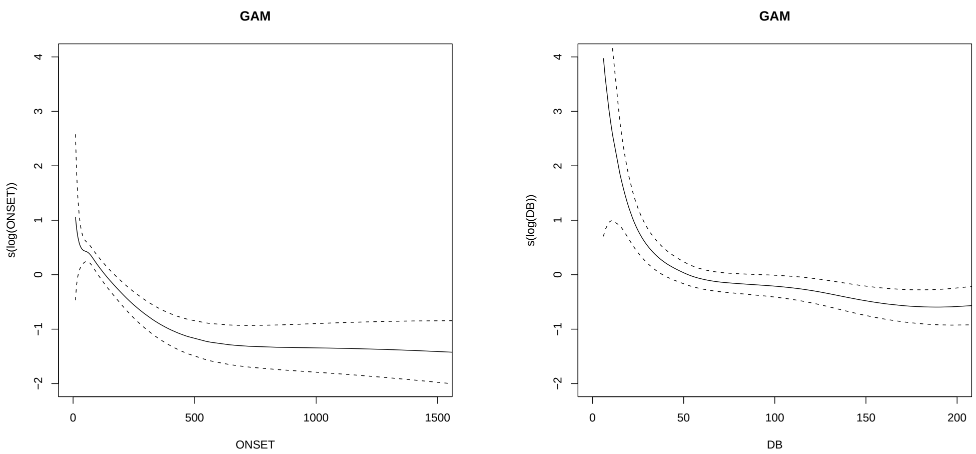

Figure 5: Estimated smoother for Symptom onset time and Door to Balloon time for GAM model of logistic regression for reperfusion efficacy (pointwise standard errors included).

5 Estimation of the Odds Ratios

We are particulary interested in studying logistic regression parameters β_j (see (1)) because they represent the change in log odds for an increase of one unit in the predictor variables; in fact let us denote by $\mathbf{X_{j+}} = (X_1, ..., X_j + 1, ..., X_p)$ we have

$$\beta_j = \text{logit}\{\text{P}(\mathbf{X_{j+}})\} - \text{logit}\{\text{P}(\mathbf{X})\} = \log\left(\frac{\text{P}(\mathbf{X_{j+}})/(\mathbf{1} - \mathbf{P}(\mathbf{X_{j+}}))}{\text{P}(\mathbf{X})/(\mathbf{1} - \mathbf{P}(\mathbf{X}))}\right); \tag{3}$$

exponentiating β_j we obtain the odds ratio of the contingency table we can built from suitable categorization of model covariates; and odds ratios are the association parameters that clinicians most widely use to describe, evaluate and communicate the results related to their studies.

Traditionally, statistical inference for contingency tables and regression parameters has relied heavily on large-sample asymptotic results for sampling distributions of parameter estimators. If we wanted to make inference about them, we would need their distribution to satisfy some asymptotic properties (see Agresti 1992), but the number and features of our data do not support them. On the other hand, a wide literature exists on exact inference for small samples in unbalanced contingency tables (see Agresti 1999). In fact large-sample approximations can result very poor and useless when we treat contingency tables unbalanced and containing both small and large frequencies such are ones arising from our data, so we had to adopt suitable techniques for exact inference.

We focused our attention on three different techniques of building exact confidence intervals for the odds ratio: the first one results from inverting the Fisher

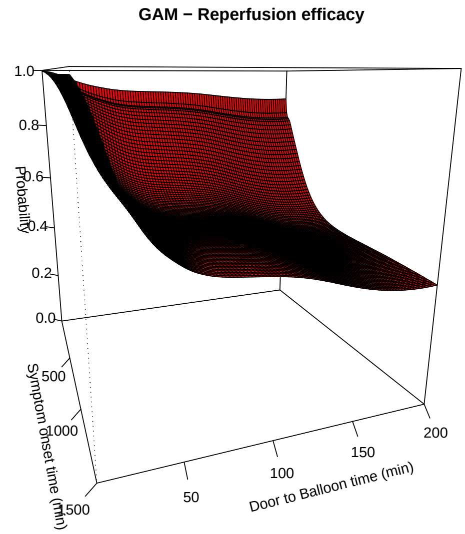

Figure 6: Reperfusion efficacy surface estimated by GAM of logistic regression.

exact test, the second one inverts an approximation of the Fisher test proposed in Gart 1962 and recommended whenever the dimension of the sample is greater than the product of the numbers of failures, and the third one inverts an unconditional test introduced in Agresti and Min 2002. To compare the performances of the three different intervals we performed a wide batch of simulations for different choices of the characterizing parameters of the contingency tables (odds ratio) and for the considered sample sizes; in all the possible cases the nominal level is guaranteed and both the Gart interval and the unconditional one tend to be shorter than the traditional Fisher confidence interval. More technical details about the numerical comparison study on these confidence intervals can be found in Ieva 2008. Tables 1 and 2 summarize, in terms of 2×2 contingency table, our data in terms of the in hospital survival and reperfusion efficacy as outcome and categorized time variables (total ischemic time and age).

		Death	Alive
	(0,150]	5	150
O	(150,300]	10	145
B	(300,600]	3	57
	> 600	4	27

		Death	Alive
A	(0,65]	6	216
g	(65,80]	7	144
e	(80,100]	9	30

Table 1: Contingency tables of survival versus Onset to Balloon time and Age.

Figures 7 and 8 show both the Gart and unconditional confidence intervals for Tables 1 and 2, respectively.

		Reduction $< 70\%$	Reduction $> 70\%$
	(0,150]	16	116
O	(150,300]	22	107
B	(300,600]	12	33
	> 600	12	14

Table 2: Contingency tables of efficacy versus Onset to Balloon time.

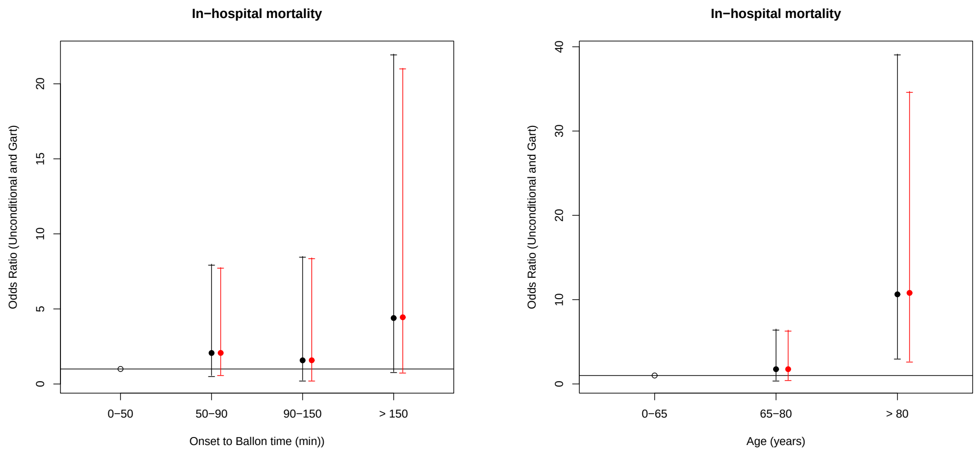

Figure 7: 95% confidence intervals for the odds ratio summarizing the relationship between in-hospital mortality and Onset to Balloon time and age: unconditional confidence interval (black), Gart confidence interval (red).

6 Conclusions and further developments

This case study indicates the efficacy of an emergency network in a complex urban area, and a simple instrument such as pre-hospital electrocardiogram teletrasmission has been proved to be one of the most critical factor necessary to keep the Door to Balloon time within the suggested limits also in a urban environment where most of STEMI patients are treated with PCI. In our data we find a critical association between pre-hospital and intra-hospital times and the principal outcomes: in-hospital survival and reperfusion efficacy. This pioneering study in such a complex urban reality supported the effort made by the 118 Dispatch Center to give the possibility of implementing electrocardiogram teletrasmission also in basic rescue units and also it will be the starting point in the design of the statistical study, part of the strategical program of the Lombardia region. The complexity of the data collection of this forthcoming register of STEMI patients in the region, will impose us the study of suitable new statistical tools in order to handle also, for example, overdispersion effects due to the wide variability in the data sources.

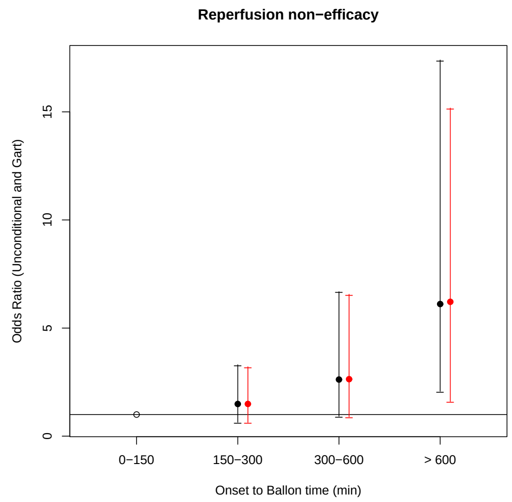

Figure 8: 95% confidence intervals for the odds ratio summarizing the relationship between reperfusion non efficacy and Onset to Balloon time and age: unconditional confidence interval (black), Gart confidence interval (red).

References

[1] A.Agresti: *A survey of Exact Inference for Contingency Tables*, Statistical Science (1992), **7**,1: 131-153

[2] A.Agresti: *On Logit Confidence Intervals for the Odds Ratio with Small Samples*, Biometrics (1999), **55**, 597-602

[3] A.Agresti, Y.Min: *Unconditional small-sample confidence intervals for the odds ratio*, Biostatistics (2002), **3**, 3, 379-386

[4] L.Breiman, J.H.Friedman, R.A.Olshen, C.J.Stone: *Classification and Regression Trees*, Wadsworth & Brooks, Monterey, California (1984)

[5] L.Breiman: *Random Forest*, Machine Learning (2001), **45**, 1: 5-32

[6] J.Gart: *Approximate Confidence Limits for the Relative Risk*, Journal of Royal Statistics Society B (1962), **24**, 2, 454-463

[7] J.Gart: *An exact test for comparing matched proportions in crossover design*, Biometriks (1969), **56**, 1, 75-80

[8] J.Gart, D.G.Thomas: *Numerical Results on Approximate Confidence Limits for the Odds Ratio*, Journal of Royal Statistics Society B (1972), **34**, 3, 441-447

[9] N.Grieco, E.Corrada, G.Sesana, G.Fontana, F.Lombardi, F.Ieva, M.Marzegalli: *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, 2008. [Online], http://mox.polimi.it/it/progetti/pubblicazioni/quaderni/10-2008.pdf

[10] T.Hastie, R.Tibshirani: *Generalized Additive Models*, Statistical Science (1986), **1**, 3: 297-318

[11] T.Hastie, R.Tibshirani: *Generalized Additive Models: some applications*, Journal of the American Statistical Association (1987),**82**, 398: 371-386

[12] T.J.Hastie, R.J.Tibshirani: *Generalized Additive Models*, Chapman & Hall/CRC, New York (1999)

[13] F. Ieva: *Modelli statistici per lo studio dei tempi di intervento nell'infarto miocardico acuto*, Master Thesis (2008), Math. Dept., Politecnico, Milano, [Online] http://mox.polimi.it/it/progetti/pubblicazioni/tesi/ieva.pdf

[14] R.L.MacNamara, Y.Wang, J.Herrin, J.P.Curtis, E.H.Bradley et al: *Effect of Door-to-Balloon Time on Mortality in Patients with ST-Segment Elevation Myocardial Infarction*, Journal of America College of Cardiology (2006), **47**, 11

[15] H.H.Ting, H.M.Krumholtz, E.H.Bradley, D.C.Cone, J.P.Curtis et al.: *Implementation and Integration of Prehospital ECGs into System of Care for Acute Coronary Sindrome*, Circulation (2008), online publication, August 13th, http://circ.ahajournals.org

[16] H.Jneid, G.C.Fonarow, C.P.Cannon, I.F.Palacios, T.Kilic et al.: *Impact of Time of Presentation on the Care and Outcomes of Acute Myocardial Infarction*, Circulation (2008), Vol.117: 2502-2509

[17] C.P.Cannon, C.M.Gibson, C.T.Lambrew, D.A.Shoultz, D.Levy, W.J.French, J.M.Gore, W.D.Weaver, W.J.Rogers, A.J.Tiefenbrunn: *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 (2000), Giugno, Vol.283, 22: 2941-2947

[18] H.H.Ting, H.M.Krumholtz, E.H.Bradley, D.C.Cone, J.P.Curtis et al.: *Implementation and Integration of Prehospital ECGs into System of Care for Acute Coronary Sindrome*, Circulation (2008), online publication, August 13th, http://circ.ahajournals.org

MOX Technical Reports, last issues

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

05/2009 F. Ieva, A.M. Paganoni:
A case study on treatment times in patients with ST-Segment Elevation Myocardial Infarction

04/2009 C. Canuto, P. Gervasio, A. Quarteroni:
Finite-Element Preconditioning of G-NI Spectral Methods

03/2009 M. D'Elia, L. Dedé, A. Quarteroni:
Reduced Basis Method for Parametrized Differential Algebraic Equations

02/2009 L. Bonaventura, C. Biotto, A. Decoene, L. Mari, E. Miglio:
A couple ecological-hydrodynamic model for the spatial distribution of sessile aquatic species in thermally forced basins

01/2009 E. Miglio, C. Sgarra:
A Finite Element Framework for Option Pricing the Bates Model

28/2008 C. D'Angelo, A. Quarteroni:
On the coupling of 1D and 3D diffusion-reaction equations. Applications to tissue perfusion problems

27/2008 A. Quarteroni:
Mathematical Models in Science and Engineering

26/2008 G. Aletti, C. May, P. Secchi:
A Central Limit Theorem, and related results, for a two-randomly reinforced urn

25/2008 D. Detomi, N. Parolini, A. Quarteroni:
Mathematics in the wind

24/2008 V. Bacchelli, A. Veneziani, S. Vessella:
Corrosion detection in a 2D domain with a polygonal boundary