METMA V International Workshop on Spatio-Temporal Modelling
Santiago de Compostela
30 Jun - 2 Jul 2010

Joint analysis of longitudinal and survival AIDS data in Brazil

Martins R.[1,4,*], Silva G.L.[2,4] and Andreozzi V.[3,4]

[1] *Centro de Investigação Interdisciplinar Egas Moniz (ciiEM), Escola Superior de Saúde Egas Moniz; ruimartins@ymail.com*
[2] *DM - Instituto Superior Técnico, Universidade Técnica de Lisboa; gsilva@math.ist.utl.pt*
[3] *DEIO, Faculdade de Ciências da Universidade de Lisboa; valeska.andreozzi@fc.ul.pt*
[4] *Centro de Estatística e Aplicações da Universidade de Lisboa*
[*] *Corresponding author*

Abstract. *Joint analysis of longitudinal and survival data has received increasing attention in the recent years, especially for AIDS data. These data have been usually analysed considering time to event data (survival outcome) or repeated measurements (longitudinal outcome) separately. As both outcomes are observed in one individual, a joint modelling of longitudinal and survival data is more appropriate because it takes into account the dependence between the two types of responses [3]. Guo and Carlin [2] addressed the problem of joint analysis by proposing a Bayesian hierarchical model, obtaining estimates for the parameters of interest via Markov chain Monte Carlo (MCMC) methods. We here employed Guo-Carlin's method for jointly modelling longitudinal and survival data for a cohort of patients with HIV/AIDS in Brazil. In addition, we also include spatial random effects to account for the unobserved heterogeneity amongst the Brazilian states. We conclude that the Bayesian joint model presents considerable improvements in the median survival time distributions when compared with those obtained through longitudinal and survival models separately.*

Keywords. *Joint model; Longitudinal data; Survival data; Bayesian inference; MCMC methods.*

1 Introduction

For most cancer and AIDS clinical trials, longitudinal and survival data have been usually analysed considering time to event data (survival outcome) or repeated measurements (longitudinal outcome) separately [1]. In that situation, as both outcomes are observed in one subject, separately modelling does not take into account the dependence between the two types of responses. In order to overcome this problem, a reasonable/powerful idea is a joint modelling of longitudinal and survival data. For instance,

Henderson, Diggle and Dobson [3] proposed a likelihood-based joint model using the EM algorithm to connect the longitudinal and survival response with a zero-mean latent bivariate Gaussian process. In addition, Guo and Carlin [2] addressed the problem of joint analysis by proposing a Bayesian hierarchical model, which may be implemented in the software WinBUGS [4], obtaining estimates for the parameters of interest via Markov chain Monte Carlo (MCMC) methods. Despite the complexity of the Bayesian approach, the joint analysis is straightforward to implement. We here employed Guo-Carlin's method for jointly modelling longitudinal and survival data for a cohort of patients with HIV/AIDS in Brazil. We make inference for several parameters of interest via Markov chain Monte Carlo (MCMC) methods. Finally, we show some results, e.g., the Bayesian joint model presents considerable improvements in the median survival time distributions when compared with those obtained through longitudinal and survival models separately.

2 Joint modelling

In order to connect the longitudinal and survival responses, Henderson, Diggle and Dobson [3] proposed a zero-mean latent bivariate Gaussian process $W_i(t) = (W_{1i}(t), W_{2i}(t))$ at time t. The repeated measurements and time to event data are assumed independent given the linking process and the covariates of interest. So, the longitudinal model described by a linear mixed effects model

$$Y_{ij} = \mu_i(t_{ij}) + W_{1i}(t_{ij}) + \varepsilon_{ij}, \tag{1}$$

incorporates fixed effects ($\mu_i = \mathbf{x}_{1i}^T(t_{ij})\boldsymbol{\beta}_1$), random effects ($W_{1i}$) and pure measurement errors ($\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)$), while the parametric or semi-parametric survival model can accommodate namely frailty terms (W_{2i}). Notice that $W_{2i} = 0$ means separately longitudinal and survival modelling.

We assumed that the longitudinal process follows equation (1) and the survival time for the i^{th} subject follows a Weibull distribution, $t_i \sim \text{Weibull}(r, \mu_i(t))$, where $r > 0$ and

$$\log(\mu_i(t)) = \mathbf{x}_{2i}^T(t)\boldsymbol{\beta}_2 + W_{2i}(t). \tag{2}$$

The event intensity at time t was given by $\lambda_i(t) = rt^{r-1}\mu_i(t) = rt^{r-1}\exp\left\{\mathbf{x}_{2i}^T(t)\boldsymbol{\beta}_2 + W_{2i}(t)\right\}$ and reduces to the exponential distribution if $r = 1$.

3 AIDS data analysis

Figure 1 compares the estimated posterior median survival time distributions for two patients in our study. The first (Patient 565) female, 28 years old, entered the study with an AIDS diagnosis and lives in south. The second (Patient 730) male, 29 years old, have an AIDS diagnosis at study entry and lives in Southwest. Both were censored at days 1361 and 1774, respectively. The bottom plots compares the posterior median survival time distributions of the joint model with the separate model for the patients. The joint result markedly differs from the separate result and significantly increasing the survival time for patients 565 and 730. Moreover, the joint model reverses the separate model findings, in the sense that the patients with the "good" CD4 trajectory are now predicted to survive much longer than the patient with the "bad" trajectory.

In short, our aim has been to study the relationship of CD4 cells (longitudinal outcome) with time to death (survival outcome) in predicting the median survival time of HIV/AIDS patients. We showed that

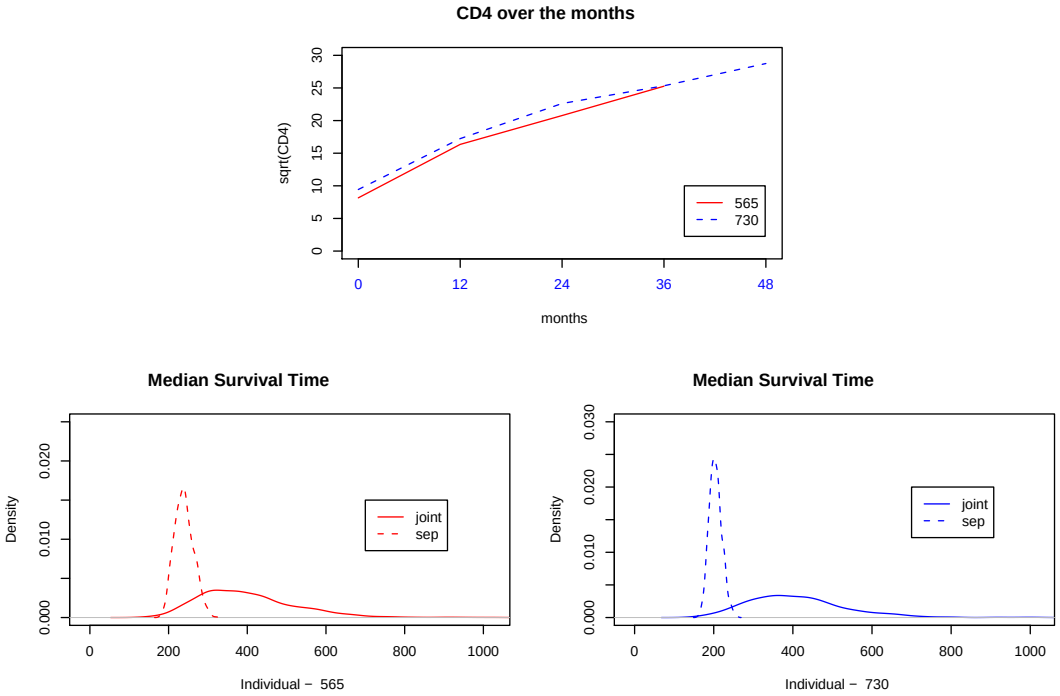

Figure 1: Observed data and estimated posteriors of median survival time for patients 565 and 730.

the Bayesian joint model presents considerable improvements in the median survival time distributions when compared with those obtained through longitudinal and survival models separately. We intend to extend that analysis incorporating spatial random effects to account for the unobserved heterogeneity amongst the Brazilian states.

Acknowledgments. The authors would like to thank Maria Goretti Pereira Fonseca Medeiros and Cláudia Medina Coeli for providing the data and MEGA project: PTDC/MAT/64353/2006 for funding Valeska Andreozzi.

References

[1] Fitzmaurice, G., Laird, N. and Ware, J. (2004). *Applied Longitudinal Data Analysis*. Wiley, New York.

[2] Guo, X., and Carlin, B. (2004). Separate and joint modelling of longitudinal and event time data using standard computer packages. *The American Statistician* **58**, 16–24.

[3] Henderson, R., Diggle, P., and Dobson, A. (2000). Joint modelling of longitudinal measurements and event time data. *Biostatistics* **1**, 465–480.

[4] Lunn, D.J., Thomas, A., Best, N., and Spiegelhalter, D. (2000) WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing* **10**, 325–337.