Multivariate Longitudinal Dynamic Regression Model of eGFR and ACR for Predicting Decline in Renal Function

Inês Sousa

Department of Mathematics and Applications, Minho University (Portugal) and
School of Health and Medicine, Lancaster University (UK); isousa@math.uminho.pt

Abstract. *We will look at the problem of estimate progression of kidney disease. The most common test gives us individual urinary albumin creatinine ratio (uACR). However, doctor would also like to see the estimated glomerular flitration rate (eGFR) test in the nacional health service. The aim of this study was to investigate the rate of progression of chronic kidney diseas, according to their eGFR and ACR. We propose a multivariate analysis as an extencion of a longitudinal dynamic regression model propose in [1].*

Keywords. *longitudinal data, multivariate, kidney failure*

1 Introduction

The global epidemic of chronic kidney disease (CKD) is a significant public health issue affecting adult population. Until recently the chief methods of screening for CKD in the UK population had involved annual assessment of urine albumin creatinine ration (uACR) and serum creatinine levels. National Institute for Clinical Excellence guidelines published in September 2008 recommend annual assessment of eGFR and uACR in line with the United States National Kidney Foundation-Kidney Disease Outcomes Quality Initiative guidelines. It is recommended annual testing of these parameters regardless of their degree of urine albumin excretion to aid early detection and prevention of progression in patients with early kidney disease. The rate of progression of CKD in single studies is not well described. Decline in eGFR varies considerably between individuals. The aim of this study was to assess progression of CKD according to their eGFR and presence of albuminuria.

2 The Statistical Problem

In this problematic we consider that there are two stochastic processes measuring differently the same true process of the progression of kidney function, wich we want to make inference about. These being the eGFR and ACR processes. In previous work ([2]) only the baseline ACR is consider and this is treated as one more categorical explanatory variable. When using the longitudinal dynamic model proposed in [1], we are allowing for the different sources of variability in the data, and at the same time, we model the progression as an integrated random walk. By extending it to a multivariate contex we use information from the two eGFR and ACR processes to estimate and make predictions about the kidney function.

3 The Model

We consider a multivariate dynamic model to the longitudinal data, incorporating variability within and between patients, considering repeated measurements from same subjects. Consider (Y_{ij}^1, Y_{ij}^2) the logarithm transformation of eGFR and ACR, respectively on subject $i = 1,...,m$ at time point t_{ij}, $j = 1,...,n_i$. Where time t_{ij} here is age of subject i at time t_{ij}. We then assume that

$$Y_{ij}^1 = \mu_{ij}^1 + U_i + C_i(t_{ij}) + Z_{ij}^1$$

and

$$Y_{ij}^2 = \mu_{ij}^2 + \alpha_1 U_i + \alpha_2 C_i(t_{ij}) + Z_{ij}^2$$

The rate of progression varies randomly, both between subjects and within subjects over time. The formal specification of the model for the time sequence (Y_{ij}^1, Y_{ij}^2), is conditional on a subject specific U_i and time specific random effect $C_i(t_{ij})$. The latter is the dynamic part of the model. Moreover, the latter is the true kidney function, that is shared by the two eGFR and ACR. Notice that, we might need two constants in this terms due to the different scales of the responses.

The Z_{ij}^1 and Z_{ij}^2 are mutually independent measurement errors, $Z_{ij}^1 \sim N(0, \tau_1^2)$ and $Z_{ij}^2 \sim N(0, \tau_2^2)$. The subject specific random effects U_i are mutually independent realisations of the distribution $U_i \sim N(0, \omega^2)$. Finally, as we are interested in the rate, we consider the integral of the Brownian motion, $C_i(t_{ij}) = \int_0^{t_{ij}} B_i(u)du$. Where, $B_i(t_{ij})$ is a subject time specific rate of change which evolves over time as a Brownian motion with starting value $B_i(t_0 = 0) = 0$, $\text{Var}(B_i(t)) = \sigma^2 t$ and $\text{Cov}\{B_i(t), B_i(s)\} = \sigma^2 \min\{t, s\}$. Therefore, what we are interested in is the subject time specific rate of change, here the process $B_i(t_{ij})$. We will be calling the model as the Integrated Random Walk (IRW) model. As we are imposing the condition of $B(t_0 = 0) = 0$ consequently also $C(t_0 = 0) = 0$. Notice that, $B(\cdot)$ is a Markov process, where

$$B_i(t)|B_i(s) \sim N\left(B_i(s), (t-s)\sigma^2\right).$$

Acknowledgments. The author would like to thank CMAT from Minho University.

References

[1] Diggle, P.J. and Sousa, I. (2010). Real-Time Monitoring Progression Towards Renal Failure in Primary Care Patients. *under submission*

[2] Hoefield, R., Kalra, P., Baker, P., Sousa, I., Diggle, P., Gibson, M., O'Donoghue, D., Middleton, R., New, J. (2010). The use of eGFR and ACR to predict decline in renal function in people with diabetes. *under submission*