A Generalized Autoregressive Conditional Heteroscedasticity Model (GARCH) to Analyze Continuous Blood Glucose Monitoring Data for Diabetic Patients

Sanjoy K. Paul and Rury R. Holman

The development of continuous blood glucose monitoring systems (CGMS) has made new it possible to capture ambulatory diurnal glucose profiles to assess glycaemic variability in Type 1 & Type 2 diabetic patients. CGMS can generate glucose time series data every 1 – 5 minutes for up to three days. The information obtained from CGMS profiles has great potential to help clinicians in terms of understanding the dynamics of glucose fluctuations, making inferences about the likelihood of hypoglycaemia and hyperglycaemic episodes and to evolve the right insulin management regimen for specific patients. Extracting robust clinical inferences from CGMS data can only be achieved if the extensive "time based" chaotic information contained within these glucose profiles can be explored using valid mathematical methods. CGMS has only recently become routinely available and the mathematical methodologies advocated so far to analyse the glucose profiles obtained are far from meeting the clinical needs. We have investigated whether a more comprehensive approach to analyzing CGMS profiles can produce new opportunities to improve diabetes management.

Modelling time-dependent CGMS data is not straightforward, especially when the glycaemic variations for a particular patient are highly sensitive to instantaneous physiological and clinical changes[1]. Our recent experience of dealing with the CGMS profiles arising from an ongoing clinical trial suggests the failure of the proposed stochastic volatility models[2] and the non-linear state-space models[3] to forecast the glucose realization even for very short time period. One of the reasons for the failure of the proposed methods to capture the glucose fluctuations is the lack of realization of the fact that there exists an effect of the inherent insulin dynamics in human body which not only effects the glucose level at present time point, but it also have a past / future lag effect on the glucose realization. This means it is important to capture the effects of inherent insulin dynamics to address the volatility in the glucose profile. We propose a GARCH model with conditional mean and variance of the type

$$y_t = x_t\beta + \varepsilon_t, \varepsilon_t \sim N(0, \sigma_t^2), \quad \sigma_t^2 = \tau_0 + \tau_1\varepsilon_{t-1}^2 + ... + \tau_m\varepsilon_{t-m}^2 + \partial_1\varepsilon_{t-1}^2 + ... + \partial_k\varepsilon_{t-k}^2;$$

where τ & δ are ARCH & GARCH parameters respectively with ε_t^2 as the innovations. To allow the conditional variance of the series to influence conditional mean glucose, where conditional variance can enter the mean equation through non-linear transformation function g(), we have the model:

$$y_t = x_t\beta + \sum_{i=1}^{p} \psi_i g(\sigma_{t-i}^2) + \sum_{j=1}^{p} \rho_j \{y_{t-j} - x_{t-j}\beta - \sum_{i=1}^{p} \psi_i g(\sigma_{t-i-j}^2)\} + \sum_{k=1}^{q} \theta_k \varepsilon_{t-k} + \varepsilon_t;$$

where β, ψ, ρ and θ are regression, ARCH, autoregressive (AR) and moving average (MA) parameters respectively with g being the non-linear function. We compared the ARCH and GARCH models with AR process of order 1 & 2. The GARCH approach addresses the near-future glucose forecasting issue more reliably compared to the other proposed state-space and stochastic methods.

[1] Paul S et al. (2007) : Modelling CGM Data – A Statistical Challenge. *ISCB 07 Abstract Book:*117.
[2] Kovatchev et al. (2005): *Diabetes technology & Therapeutics 2005: 7(6)*: 849 – 861.
[3] Magni et al. (2006): *IEEE Transactions on Biomedical Engineering 2006: 53(6)*: 977- 985.