A Decade of Joint Mean-Covariance Modelling: What has the Industry learned?

Gilbert MacKenzie[1]

[1] Centre of Biostatistics, Department of Mathematics & Statistics, The University of Limerick, Limerick, Ireland and ENSAI, Bruz, France

Abstract

The conventional approach to modelling longitudinal RCT data places considerable emphasis on estimation of the mean structure and much less on the covariance structure, between repeated measurements on the same subject. Often, the covariance structure is thought to be a 'nuisance parameter' or at least not to be of primary 'scientific interest' and little effort is expended on elucidating its structure. In particular, the idea that intervention might affect the covariance structure rather than, or as well as, the mean rarely intrudes.

A decade on, we shall argue that these ideas are rather passé and that from an inferential standpoint the problem is symmetrical in both parameters μ and Σ. Throughout, we will distinguish carefully between joint estimation which is now relatively routine and joint model selection which is not.

At first sight the task of estimating the structure of Σ, from the data, rather than from a prespecified menu, may seem daunting, whence the idea of searching the entire covariance model space, $\{\mathcal{C}\}$, for Σ, may seem prohibitive. Thus, the final demand that we conduct a simultaneous search of the Cartesian product of the mean-covariance model space, $\{\mathcal{M} \text{ x } \mathcal{C}\}$, may seem impossible. However, below, we shall accomplish all three tasks elegantly for a particular, but very general, class of covariance structures, $\{\mathcal{C}^*\}$, defined below.

The technique is based on a modified Cholesky decomposition of the usual marginal covariance matrix $\Sigma(t, \theta)$, where t represents time and θ is a low-dimensional vector of parameters describing dependence on time. The decomposition leads to a reparametrization, $\Sigma(t, \varsigma, \phi)$, in which the new parameters have an obvious statistical interpretation in terms of the natural logarithms of the innovation variances, ς, and autoregressive coefficients, ϕ. These unconstrained parameters are modelled, parsimoniously, as different polynomial functions of time.

In this talk we trace the history of the development of joint mean-covariance modelling over the last decade, from Pourahmadi's seminal paper in 1999 to recent times, discuss current research in this paradigm and remark on the lack of impact on trial design and analysis.

Key References

Pan J. X. and MacKenzie G (2003). On modelling mean-covariance structures in longitudinal studies. *Biometrika*, **90**, 239-244.

Pourahmadi, M. (1999). Joint mean-covariance models with applications to longitudinal data: Unconstrained parameterisation. *Biometrika* **86**, 677-90.