CONTROL ID: 3343284
CONTACT (INSTITUTION ONLY): Leiden University Medical Center
CONTACT (NAME ONLY): Mirko Signorelli
TITLE: How to predict a survival outcome using longitudinal and high-dimensional omics data
ABSTRACT BODY:
Abstract Body: Longitudinal and high-dimensional measurements have become increasingly common in biomedical research. In survival analysis, longitudinal covariates are usually accounted for using joint models for longitudinal and time-to-event outcomes, while penalized survival models are employed to tackle high-dimensional sets of predictors. However, methods to predict survival outcomes using covariates that are both longitudinal AND high-dimensional are currently missing. In this presentation we propose a strategy to overcome this limitation.
Our work is motivated by data from the MarkMD study, which involved 157 patients affected by Duchenne Muscular Dystrophy (DMD), a rare genetic disorder that leads to loss of ambulation (LoA) and premature death. No cure for DMD is available yet and non-invasive markers are urgently needed to monitor disease progression and predict disease milestones. In the MarkMD study, an antibody-based bead array was used to measure the serum levels of 118 different proteins with 240 antibodies in a total of 303 serum samples. The aim of the study was to understand if longitudinal protein data could be used to improve the prediction of the age at which LoA occurs. Important features of these data are 1) the longitudinal nature of the predictors, 2) the presence of strong correlations between antibodies that target the same protein and 3) high-dimensionality.
To tackle this prediction problem, we propose a modelling approach whereby the longitudinal trajectories of antibodies measuring the same protein are modelled using latent process mixed models, a multivariate extension of mixed models that allows us to deal with problems (1) and (2). Based on these models, we summarize the longitudinal information via the predicted protein- and antibody-specific random effects. Then, we employ the (high-dimensional) summaries thus derived as predictors in an elastic-net penalized Cox model, from which we derive predictions of time to LoA. We first compare the predictive performance of our model to that of simpler prediction strategies using simulated data, considering different scenarios that allow to understand how our approach can improve predictions. Then we apply our model to the MarkMD data, showing that predictions of time to LoA in DMD can be considerably improved by fully exploiting the available information on the dynamic evolution of a large number of biomarkers.
AUTHORS/INSTITUTIONS: M. Signorelli, R. Tsonaka, Biomedical Data Sciences, Leiden University Medical Center, Leiden, NETHERLANDS|P. Spitali, Human Genetics, Leiden University Medical Center, Leiden, NETHERLANDS|