In life sciences, there has been an increasing trend towards analyses involving not only one single endpoint as the primary focus of the analysis but focusing rather on the simultaneous analysis of multiple outcomes. One common example are time-to-event and longitudinal data where, for the same set of individuals, one or multiple markers of interest have been observed repeatedly over time (longitudinal data), in conjunction with observations on the time to a certain event (time-to-event data). Separate analyses of such data will typically be inappropriate when one is interested in the relationship between the longitudinal outcomes and the time-to-event data and may lead to biased estimates as well as a loss of efficiency. Joint models overcome these difficulties by simultaneously analyzing the longitudinal outcome variables with the duration process. In such a context, a wide variety of approaches have been developed, including

• shared random effect models that link the longitudinal and the time-to-event model by a shared set of random effects that induce marginal correlation between the models,

• shared predictor models that generalize this idea by considering a more flexible subset of the predictor of the longitudinal and the time-to-event model that is shared by both model components,

• pattern-mixture models that factor the joint distribution of longitudinal and time-to-event data into the marginal distribution of the time-to-event data (possibly given random effects) and the conditional distribution of the longitudinal data given the time-to-event data, and

• selection models that are based on the same factorization idea as pattern mixture models but reverse the order, i.e. they are based on the marginal distribution of the longitudinal data and the conditional distribution of the time-to-event data given the longitudinal data.

All approaches have in common that the main objective is to provide a framework for the simultaneous analysis of the longitudinal outcomes and the time-to-event data. In particular, joint modelling approaches aim at characterizing the joint distribution of the longitudinal outcomes and the time-to-event data in different ways depending on the framework to avoid the bias and loss of efficiency that can appear in separate treatments.

While the simultaneous analysis of longitudinal and time-to-event data is the most common example of “joint modelling”, there are other forms as well, comprising, for example, the simultaneous analysis of multiple markers in a longitudinal study based on multivariate response models. For this special issue, we interpreted the term “joint models” in a broad sense such that it also includes alternative forms of joint modelling (see also the summaries of the contributions below).
This special issue was initiated early in 2016 preceding the “Second Galician-Portuguese meeting of Biometry, with applications to Health Sciences, Ecology and Environmental Sciences” (BIOAPP-2016, http://biometria.sgapeio.es/) which took place in Santiago de Compostela, Spain, from June 30 to July 2, 2016 and the workshop “Joint Modeling and Beyond 2016” (http://www.uhasselt.be/Joint-Modeling-and-Beyond) that took place at the University of Hasselt, Belgium, from April 14 to April 15, 2016. The Second Galician-Portuguese Meeting of Biometry was organized jointly by the Portuguese Statistical Society and the Galician Society for the Advancement of Statistics and Operations Research with the aim of disseminating the latest advances in the development and application of statistical and mathematical methods in biology, medicine, ecology, psychology, pharmacology, agriculture, environment and other health and life sciences. The Joint Modelling and Beyond workshop was conducted under the IAP-StuDyS (Belgian InterUniversity Attraction Pole Network IUAP-P7/06) and was jointly organized by I-BioStat in Hasselt and Leuven, the University Medical Centre of Rotterdam, the Interdisciplinary Centre of the University of Santiago de Compostela and the Biostatnet Network.

Of course, although participants of these two workshops where particularly encouraged to submit papers that fit into the scope of this special issue, the possibility to submit was not restricted as is also apparent from the contributions collected in this special issue. Some of these contributions focus on developing and evaluating novel inferential procedures for joint models.

Waldmann et al. (2017) provide a useful alternative to commonly used maximum likelihood and Bayesian methods, employing boosting methodology that allows the analyst to simultaneously estimate predictors and select influential variables. The method is designed to work well with high-dimensional data as well. In other words, the method introduces machine-learning methodology in the joint modelling world.

Do Ha et al. (2017) utilize the framework of h-likelihoods for the estimation of joint models for longitudinal and time-to-event data. They derive a set of joint iterative least squares equations that lead to simultaneous estimation of the model parameters and of the random effects / frailty terms. Several extensions of standard joint models are considered, including risks and multiple random effects. They empirically investigate the performance of their fitting procedure, and they also compare it with the more traditional approach that entails numerical integration for calculating the marginal log-likelihood. This procedure is implemented in the function jmfit() from the R package frailtyHL.

Bayesian methodology for joint models is another area addressed by some contributions. Köhler et al. (2017) develop a generic framework of flexible additive joint models that encompasses the specification of a variety of effects, such as smooth nonlinear, time-varying and random effects, in the longitudinal and time-to-event parts of the models. An implementation is provided in the R-package bamlss. Their extensions are motivated by the investigation of the relationship between disease-specific markers, namely autoantibodies, and the progression of type 1 diabetes. High nonlinearities can be captured in these models using Bayesian penalized splines which allows for new insights into the disease progression.

Martins et al. (2017) propose a Bayesian hierarchical joint model for the analysis of disease progression for HIV/AIDS patients collected in all the 27 states of Brazil during the period 2002–2006. In the longitudinal part, they utilize penalized splines to identify flexible
trends in the CD4 profiles of the patients while the time-to-event part combines a spatially clustered specification with a cure rate-type model to allow for a fraction of long-term survivors.

Rue et al. (2017) study patient-ventilator asynchronies (PVAs) which occur when timing of the ventilator cycle is not simultaneously timed with the patient respiratory cycle. Compared to most previous analyses, the authors add the PVA data in terms of an index of asynchronies (AI) and the sequential organ failure assessment (SOFA) score to assess overall severity. For the first, a Bayesian joint model of bivariate longitudinal and competing risks data is introduced. The longitudinal process includes a mixed effects model for the SOFA score and a mixed effects beta regression for the AI. The authors find that the SOFA score is strongly related to vital status but only slight evidence is given that PVAs provide a more accurate indication of death prognosis.

Another branch deals with joint models involving multivariate longitudinal data.

Guler et al. (2017) present a two-stage approach for estimating joint models with multiple longitudinal outcomes and a time-to-event response. Their approach is an extension of the pairwise fitting procedure of Fieuws and Verbeke (2006) and provides valid inferences for exogenous time-varying covariates. The authors illustrate their approach using numerical studies and a case study from nephrology research. The authors provide also SAS code to implement their approach.

Nassiri et al. (2017) propose fast precision estimation of parameters in high-dimensional multivariate joint models. Such precision estimation is based on the multiple outputation (MO, Hoffman et al., 2001; Follmann et al., 2003) and it can be easily implemented in a pairwise approach (Fieuws and Verbeke, 2006; Fieuws et al., 2007). The performance of the MO-based correction was compared with that obtained using the sandwich correction, through simulations and a real data analysis. The results obtained suggest that the MO-based correction is considerably faster and therefore constitutes an easy and efficient way to estimate the precision when dealing with high-dimensional multivariate joint models.

Duarte et al. (2017) present a bivariate model, relevant for the study of breast cancer risk studied as a function of age at menarche, age at menopause, as well as cohort effects and spatial trends. Not only are the two outcomes (age at menarche and lifespan cycle) of interest, so is the association between both. To study this association, they propose a Bayesian multivariate structured additive distributional regression model.

Another important topic in joint modelling concerns prediction.

Sweeting (2017) illustrates how dynamic predictions obtained from a joint longitudinal and survival model can be used to inform an optimal intervention time for individuals under surveillance in an abdominal aortic aneurysm (AAA) screening programme. The joint modelling approach is proposed to associate longitudinal measurements of aortic diameter with the risk of aneurysm rupture. The results indicate that the decision to intervene, currently based on a diameter value greater than a threshold of 5.5 mm, could be made more personalised and dynamic in a decision modelling approach. Using AAA data together with external information on operative risk, some initial recommendations regarding the optimal intervention time and the diameter threshold for elective surgery can be made in the context of the AAA screening programme.

Rizopoulos et al. (2017) present and compare two popular statistical techniques, namely landmark analysis and joint models for longitudinal and time-to-event data, to provide dynamically updated estimates of survival probabilities with time-dependent variables. To assess the quality of the derived predictions from the two approaches, different measures
of discrimination and calibration are considered after adjustment to the context of longitudinal biomarkers. A simulation study and the analysis of the Aortic Valve data set suggest that, in general, there is a gain from considering the joint modeling approach instead of landmarking.

Suresh et al. (2017) focus on dynamic prediction, i.e., the use of time-dependent marker information, corrected during the course of a trial, in order to improve prediction of survival. For this to be valid, ordinarily a joint model of the marker and the survival time is needed, but it leads to considerable computational burden. The authors compare joint modeling with landmarking, an approximate but computationally efficient alternative.

In summary, this special issue summarizes some recent advances in joint modelling (in a broad sense) that address challenging methodological and applied questions and therefore provide a good overview both on the current state of the art in joint modelling and promising trends for future developments.

References


Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Suarez, CC; Klein, N; Kneib, T; Molenberghs, G; Rizopoulos, D

Title:
Editorial "Joint modeling of longitudinal and time-to-event data and beyond"

Date:
2017-11-01

Citation:

Persistent Link:
http://hdl.handle.net/11343/293806