Summary of the Current Work

My research extends the standard approach of Faucett and Thomas [3] and Wulfson and Tsiatis [8] in joint modeling in several ways. In particular, I developed a two-stage procedure that can be used as an alternative to joint modeling approach in case when the longitudinal measurements are collected before the start of follow-up for survival response [7]. In contrast with the standard joint modeling setting, the longitudinal responses do not constitute an endogenous time dependent variable measured at the same period as the time to event. Nevertheless, the problem of measurement error still remains. The proposed two-stage procedure handles the problem of measurement error via Monte Carlo sampling from the posterior distribution of the random effects. I have applied this approach for nonlinear longitudinal response and compared the results with the "naive" plug-in approach when the uncertainty about the estimates from the first step is not taken into account, as well as with the full Bayesian approach.

In addition I have considered categorical longitudinal responses in the presence of competing risks. We showed how this problem can be handled using multi-state models techniques. In particular we used the pseudo-values approach introduced by Andersen et al. [2] and applied it for the Aalen-Johansen estimator of the state occupation probabilities [9]. To address the problem of those competing events a multinomial approach was used for the next state given the previous state observed. This has a great advantage compared to other methods for non-Markov models where the history of the process is of interest and no standard approaches are available. Moreover we formulated the problem in the joint modeling framework and proposed a Bayesian model for joint modeling of categorical longitudinal data and time-to-event response taking into account the presence of competing risks [8].

In addition for the developed Bayesian joint model I have derived posterior predictive distributions for the longitudinal and event time outcomes [8]. Additionally, we have also examined the impact of different parameterizations of the joint model on the obtained predictions. Further we have compared the joint modeling technique for making dynamic predictions with an older method for producing such predictions, called landmarking [10]. We showed how survival probabilities are obtained under each method and what are the differences in the underlying assumptions. The current work is focused on the functional relationship between the two processes which affects predictions. To assess the quality of the derived predictions from the two approaches we work on different measures of discrimination and calibration, suitably adjusted to the context of longitudinal biomarkers.

Directions for Future Work

The further work is to consider multiple correlated biomarkers since majority of settings in joint modeling literature present models only with a single longitudinal outcome and time-to-event. The separate analysis per longitudinal outcome was shown to be less efficient than a joint analysis of all the markers simultaneously ([4],[6]). In practice extending the joint model from single to multiple continuous longitudinal outcome creates similar problems as considering more categories in a model with a categorical outcome. This leads to highly time- and memory consuming estimation that often suffers from convergence problems. Therefore, it is clear that we need alternative, less computationally intensive methods that could be used in real data problems. Such an alternative could be the conditional score approach proposed by Tsiatis and Davidian [11] and extended by Song et al. [14] for the multivariate longitudinal data. This method is based on estimating equations and makes no distributional assumption on the underlying random effects, treating them as "nuisance". However, especially for the multivariate longitudinal data, the cognitional score approach reduces considerably the computational complexity compared to likelihood or Bayesian approaches.

With respect to the survival outcome a multivariate extension is to consider multiple failure times per subject, such as recurrent event. This type of models require an additional submodel for the recurrent events, increasing the computational complexity.

The issue of model selection in joint modeling is also under our investigation. Up to now not many solutions have been proposed. When Bayesian methods are used for the estimation, a DIC or other Bayesian criteria could be considered. However, due to the well-known limitations of such criteria, future work could focus on developing more general measures that would allow to choose the best model based on the quality of the produced predictions in terms of calibration and discrimination, regardless the estimation method. In particular, discrimination measures that could be applied in a competing risk setting using joint models are of a special interest. In the context of time-dependent ROC curves Heagerty et al. [5] proposed several definitions of cases and controls. Saha and Heagerty [13] and Zheng et al. [16] extended this definition for the competing risks setting. Depending on the particular setting we could consider different methods of classifying subjects and use similar sampling procedure as Rizopoulos [15] to estimate ROC in the joint modeling framework. This extension could be applied to the fully Bayesian model.

References

[1] Albert PS, Shih JH. (2009) On estimating the relationship between longitudinal measurments and time-to-event data using a simple two-stage procedure. *Biometrics* **66**, 983–991.

[2] Andersen, P.K. and Klein, J.P. and Rosthoj, S. (2003) Generalized linear models for correlated pseudo-observations with applications to multi-state models. *Biometrika* **90**, 15–27. [3] Faucett, C. and Thomas, D.(1996) Simultaneously modelling censored survival data and repeatedly measured covariates. *Statistics in Medicine* **15**,1460–1461.

[4] Fieuws, S. and Verbeke, G. and Maes, B. and Vanrenterghem, Y. (2008) Predicting renal graft failure using multivariate longitudinal profiles. *Biometrika* **9**, 419–431.

[5] Heagerty, P.J. (2005) Survival Model Predictive Accuracy and ROC Curves. *Biometrics* **61**, 92-105.

[6] McCulloch, C. (2008) Joint modelling of mixed outcome types using latent variables. *Statistical Methods in Medical Research* **17**, 53-73.

[7] **Murawska, M.**, Rizopoulos, D. and Lesaffre, E. (2012) A Two-Stage Joint Model for Nonlinear Longitudinal Response and a Time-to-Event with Application in Transplantation Studies, vol. in press, *J Probability and Statistics* [DOI:10.1155/2012/194194]

[8] **Murawska, M**, Rizopoulos, D. and Lesaffre, E. Dynamic Predictions for Categorical Longitudinal Responses and Event Data with Application in Transplantation Studies, *Biostatistics*, 2013. (Submitted)

[9] **Murawska, M.**, Rizopoulos D. and Lesaffre, E. Simple analysis of non-Markov models: A case study on heart transplant data, *Statistical Modeling*, 2013. (Submitted after major revision)

[10] Rizopoulos, D., **Murawska**, M., Andrinopoulou,E-R., Molenberghs, G, Takkenberg, J.J.M and Lesaffre, E. A comparison between landmarking and joint modeling for producing predictions using longitudinal outcomes, *Biometrical Journal*, 2013. (Submitted)

[11] Tsiatis, A.A. and Davidian, M. (2001) A semiparametric estimator for the proportional hazards model with longitudinal covariates measured with error. *Biometrika* **88**, 447–58.

[12] Tsiatis A, DeGruttola V, Wulfsohn M. (1995) Modeling the relationship of survival to longitudinal data measured with error: Applications to survival and CD4 counts in patients with AIDS. *Journal of the American Statistical Association* **90**, 27–37.

[13] Saha, P. and Heagerty, P.J. (2010) Time-dependent predictive accuracy in the presence of competing risks. *Biometrics* **66**, 999–1011.

[14] Song, X. and Davidian, M. and Tsiatis, A.A. (2002) An estimator for the proportional hazards model with multiple longitudinal covariates measured with error. *Biostatistics* **3**, 511–528.

[15] Rizopoulos, D. (2011) Dynamic prediction and prospective accuracy in joint models for longitudinal and time-to-event data. *Biometrics* **67**, 819–829.

[16] Zheng, Y. and Cai, T. and Jin, Y. and Feng, Z. (2012) Evaluating prognostic accuracy of biomarkers under competing risk. *Biometrics* **68**, 388–396.