Deutsche Region

Deutsche Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e.V.

Deutsche Gesellschaft für Sozialmedizin und Prävention

Workshop

Statistical Analysis of Event and Longitudinal Data

on November 19th – 20th, 2015

of the working groups “Statistical Methods in Medicine” (IBS-DR), “Statistical Methods in Epidemiology” (IBS-DR, DGEpi), “Statistical Methods in Clinical Research” (GMDS) and “Epidemiological Methods” (DGEpi, GMDS, DGSMP)

Location:

Institute of Medical Biostatistics, Epidemiology and Informatics Mainz

University Medical Center of the Johannes Gutenberg University Mainz

Obere Zahlbacher Str. 69

55131 Mainz

Schedule

Thursday, November 19th

13:30-14:00 Welcome and Registration

14:00-15:00 Invited Talk by Hélène Jacqmin-Gadda (Bordeaux): Joint models for longitudinal data and time to event

Coffee Break

Joint models and longitudinal data

15:20-15:45 Tobias Bluhmki (Ulm): Non-parametric simulation of complex time-to-event data and its application to a simple joint model

15:45-16:10 Sandra Walser (München): Longitudinal analysis of audiometric data from the *Ohrkan* cohort study

16:10-16:35 Stephanie Brandt (Ulm): Handling missing longitudinal data during childhood in the Ulm Birth Cohort study: A comparison of multiple imputation, joint models and survival analysis for the development of overweight

Coffee Break

Repeated measurements and events

17:00-17:25 Irene Schmidtmann (Mainz): Effects of irregularly measured covariates on dynamic prediction in time to event data

17:25-17:50 Verena Schneider-Lindner (Mannheim): Evaluating change in systemic inflammatory response in septic and nonseptic polytrauma patients

17:50-18:15 Katharina Ingel (Mainz): Simulating recurrent event data with a competing terminal event when failure rates depend on multiple time scales

Get together at Proviant-Magazin (starting at 19:30)

Friday, November 20th

8:30-9:00 Gemeinsame AG-Sitzung

9:00-10:00 Invited Talk by Jan Beyersmann (Ulm): Florence Nightingale, William Farr and Competing Risks

Illness-death models

10:00-10:25 Nadine Binder (Freiburg): Estimating hazard ratios in cohort data with missing disease information due to death

10:25-10:50 Ralph Brinks (Düsseldorf): Differential mortality explains maximum of age-specific seroprevalence in Hepatitis A infections

10:50-11:20 Coffee Break

Beyond the Cox model

11:20 - 11:45 Ulrike Pötschger (Wien): Using pseudo-values for comparing long-term survival after stem-cell transplantation (SCT) with long-term survival after chemotherapy

11:45 - 12:10 Daniela Zöller (Mainz): Identifying relevant variables by simultaneous variable selection between dynamic prediction models based on pseudo-values

12:10 - 12:35 Geraldine Rauch (Heidelberg): The average hazard ratio - a good effect measure for time-to-event endpoints when the proportional hazard assumption is violated?

12:35 - 13:00 Marvin Wright (Lübeck): Random forests for survival analysis using maximally selected rank statistics

13:00-13:15 Closing

Non-parametric simulation of complex time-to-event data and its application to a simple joint model

Tobias Bluhmki[1], Claudia Schmoor[2], Arthur Allignol[1], and Jan Beyersmann[1],

[1]Institute of Statistics, Ulm University, Helmholtzstrasse 20, 89081 Ulm, Germany
email: tobias.bluhmki@uni-ulm.de, arthur.allignol@uni-ulm.de, jan.beyersmann@uni-ulm.de
[2]Studienzentrum Universtiaetsklinikum Freiburg, Elsaesser Strasse 2, 79110 Freiburg
email: claudia.schmoor@uniklinik-freiburg.de

In many practical situations multistate models can be applied as joint models for both time-dependent covariates with finite range and time-to-event data. A common starting point for simulation studies is to force hazards to follow a pre-specified parametric model fitting the original data. In contrast, the simulation procedure presented in this talk generalizes the non-parametric approach originally suggested in [1] to more complex multistate models. The hazards are derived from the increments of the estimated cumulative hazards arising from the original data and are employed for a simulation algorithm using a nested sequence of competing risk events. The present approach accounts for the time-inhomogeneous Markov assumption and allows for right-censoring, left-truncation, competing risks, and degenerated initial distributions.
The algorithm is illustrated with published data of a randomized clinical trial including leukemia patients in early or advanced disease status [2,3]. One aim was to investigate the impact of graft-versus-host disease (GVHD) prophylaxis on the time under imunosuppressive therapy (IST). Due to the time-dependent nature of IST status (back and forth switches between 'IST' and 'no IST' possible), a joint model based on an illness-death multistate model with recovery is applied [4]. Simulation checks and an overview on the current literature are also presented.

[1] Allignol A., Schumacher M., Wanner C., Drechsler C., Beyersmann J. Understanding competing risks: a simulation point of view. BMC Med Res Methodol 2011, 11:86.

[2] Finke J., Bethge W., Schmoor C., Ottinger H.D, et al. Standard graft-versus-host disease prophylaxis with or without anti-T-cell globulin in haematopoietic cell transplantation from matchd unrelated donors: a randomised, open-label, multicentre phase 3 trial. Lancet Oncol 2009, 10:855-64.

[3] Socié G, Schmoor C, Bethge WA et al. Chronic Graft-Versus-Host Disease: Long term results from a Randomized Trial on GvHD Prophylaxis with or without Anti-T-Cell Globulin ATG-Fresenius. Blood 2011, 117:6375-82.

[4] Schmoor C., Schumacher M., Finke J., Beyersmann J. Competing Risks and Multistate Models. Clin Cancer Res 2013, 19(1): 12-21.

Longitudinal analysis of audiometric data from the “Ohrkan” cohort study

Sandra M. Walser, Doris G. Gerstner, Valerie Matthäus, Alisa Weber, Christina Reiter, Stefanie Kolb, Caroline E. W. Herr

Background: Since 2009 the prospective cohort study “Ohrkan” is performed by the Bavarian Health and Food Safety Authority (LGL). The aim of this study is to estimate the exposure to leisure and work noise as well as the prevalence of audiometric notches among adolescents in Bavaria for a period of ten years. Data was collected via standardized questionnaires and clinical ear examination. Overall, 2149 of the 3846 eligible adolescents (56%) participated.
Baseline survey: Considering the total exposure to leisure noise 34% of adolescents exceeded the lower action value (LAV) of work noise limits. Among users of personal music players (PMP) 32% exceeded the LAV. In contrast, hearing of these adolescents seems not (yet) to be damaged by this noise. Among 1843 adolescents only 2.4% had an indication of noise-induced hearing loss in the audiogram.
First Follow-up: Compared to the baseline survey, a slight increase of participants exposed by the use of PMP exceeded the LAV. A considerable increase was observed in the overall burden of leisure noise.
In the recent survey, five years after the initial investigation, both a postal questionnaires and a clinical ear examination are performed. A longitudinal analysis of these data is planned to determine for the first time the long-term effects of exposure.

Title: Handling missing longitudinal data during childhood in the Ulm Birth Cohort study: A comparison of multiple imputation, joint models and survival analysis for the development of overweight

Authors: Stephanie Brandt[1], Martin Wabitsch[1], Hermann Brenner[2], Dietrich Rothenbacher[3], Jan Beyersmann[4]

[1]University Medical Center Ulm, Department of Pediatrics and Adolescent Medicine, Division of Pediatric Endocrinology and Diabetes, Ulm, Germany
[2]Cancer Research Center (DKFZ), Division of Epidemiology and Aging Research, University of Heidelberg, Germany
[3]Ulm University, Institute of Epidemiology and Medical Biometry, Ulm, Germany
[4]Ulm University, Institute of Statistics, Ulm, Germany

Background: The presence of missing data is a well-known problem in clinical studies with a longitudinal study design. One problem for statistical analyses is to deal with missing longitudinal data.

Objective: To compare three methodological approaches to deal with missing longitudinal data: (1) Multiple Imputation of missing longitudinal data and subsequent calculation of linear mixed effects model, (2) Censoring of missing data in a survival analysis, (3) Joint model with time to dropout as the survival sub-model in the real data set of the Ulm Birth Cohort Study (UBCS). The UBCS was initiated to investigate BMI trajectories during childhood and to identify early life risk factors for the development of overweight in childhood. In the UBCS some participants dropped out of the study, BMI values of these participants are missing.

Material and Methods: (a) Recruitment of study participants (UBCS): All women who were admitted to the Department of Gynaecology and Obstetris at the University of Ulm between November 2000 and November 2001 to give birth to a baby were recruited (N=1.088) (b) Study design of the UBCS: Baseline examination and regularly follow-up examinations of the cohort at child´s age of 6 weeks, 6 months, 1,2,3,4 and 6 years (BMI values (kg/m^2) has been obtained by parent and pediatricians questionnaire). Overweight in children at the respective follow-up time points: BMI $\geq 90^{th}$ age- and sex-specific percentiles of the German reference data [Kromeyer-Hauschild, 2001]. (c) Medical research question: to study if there are differences in BMI trajectories and in the risk of becoming overweight in childhood between offsprings of pre-pregnancy overweight and offsprings of pre-pregnancy normal weight women. (d) Three methodological approaches to deal with missing longitudinal data: (d1) *Multiple imputation (MI) of missing data and linear mixed effects model* (SAS 9.3 statistics software): imputation model includes gender, maternal baseline covariates, BMI values at follow-up examination, last measured BMI values, dropout and age at dropout; method: MCMC-method; number of imputation: N=5 ; linear mixed effects model: to model the BMI trajectories during childhood (0-6 years). (d2) *Survival analysis – Modeling of time to overweight* (SAS 9.3 statistics software): censoring of participants with missing longitudinal data + censoring of participants at the end of the study period, if the event did not occur. (d3) *Joint models-Modeling of time to dropout* (R Statistical Software): modeling of time to dropout as the survival sub-model. Approaches d1 and d3 are parametric and either assume Missingness at Random (MAR, d1) or allow for non MAR patterns (d3). Approach d2 is semi-parametric and makes an assumption not unlike sequential MAR. The rational to use MI was to allow for nice graphical ilustration of BMI trajectories and to investigate approach d2 also with imputed values.

Results: Each of the three methodological approaches has shown a positive association between maternal pre-pregnancy overweight and offsprings BMI values in childhood. The results of the linear mixed effects model with MI of missing longitudinal data and the joint models indicated a higher intercept at birth in offsprings of pre-pregnancy overweight than in offsprings of pre-pregnancy normal weight mothers. Both methodological approaches estimated a lower decrease in BMI values after 0.75 up to 6 years of age in offsprings of pre-pregnancy overweight than in offsprings of pre-pregnancy normal weight mothers. The survival analysis estimated a 46% higher risk of becoming overweight in childhood for offsprings of pre-pregnancy overweight mothers than for offsprings of pre-pregnancy normal weight mothers (HR_1: 1.46; 95%CI 1.22-1.68). In addition, a second survival analysis that used multiple imputed longitudinal data of BMI values has been calculated. In the data set with multiple imputed BMI data, a HR of 1.55 (HR_2:1.55, 95% CI: 1.18-2.05) has been calculated for offsprings of pre-pregnancy overweight mothers in comparison to offsprings of pre-pregnancy normal weight mothers. The two calculated HR (HR_1: 1.46; 95%CI 1.22-1.68; HR_2:1.55, 95% CI: 1.18-2.05) are close together.

Discussion: The interpretation of results of all methodological approaches to handle missing longitudinal data were comparable with each other and hence appear to robust against the various assumptions made. From a clinical point of view, the effect sizes of the survival model were easier to interpret, but required BMI thresholds for overweight. .

Effects of irregularly measured covariates on dynamic prediction in time to event data

I. Schmidtmann (1), A. Weinmann (2), A. Schritz (1, 3), D. Zöller(1), H. Binder (1)

(1) Institut für Medizinische Biometrie, Epidemiologie und Informatik and (2) I. Medizinische Klinik, Universitätsmedizin der Johannes Gutenberg-Universität Mainz (3) Hochschule Koblenz, Rhein-Ahr-Campus

Data from clinical cancer registries provide the opportunity to assess the prognosis of cancer patients during follow-up given current (and previous) health status and treatments. This can be achieved by fitting dynamic prediction proportional hazard regression models for the conditional survival.

However, in contrast to clinical studies where regular follow-up visits are scheduled according to a study protocol, in an observational setting the frequency of follow-up visits is not only driven by guidelines but also by individual needs and treating physicians' judgement. Hence the updates of laboratory values and other time varying covariates may differ between patients depending on current health status or imminent treatment change.

The naïve approach in dynamic prediction is to use the most recent laboratory values as covariates. However, this may lead to biased estimates of coefficients in regression models. We demonstrate this in a simulation study for Cox proportional hazards models.

Here, we investigate several bias correction approaches. Simple approaches are to to use the time since the last update as covariate or the number of updates since the last landmark. Another is to fit an Andersen-Gill model to the frequency of follow-up visits using baseline covariates and to include the martingale residuals as a measure of unexplained heterogeneity in patients in the main proportional hazard regression model. Alternatively, we consider frailty estimates from Andersen-Gill models for adjustment.

We illustrate the differences in estimated regression coefficients using data from the Mainz hepatocellular carcinoma registry, thereby demonstrating the potential extent of bias. We also systematically investigate bias and the effect of correction approaches in a simulation study.

Evaluating change in systemic inflammatory response in septic and nonseptic polytrauma patients

Holger A. Lindner, PhD[1], Ümniye Balaban, MSc[1,2], Timo Sturm, MD[1], Christel Weiß, PhD[2], Manfred Thiel, MD[1], Verena Schneider-Lindner, MD PhD[1]

[1]Department of Anaesthesiology and Surgical Intensive Care Medicine, Medical Faculty Mannheim, Heidelberg University, Germany

[2]Department of Medical Statistics, Medical Faculty Mannheim, Heidelberg University, Germany

Background: Sepsis is a life-threatening infectious complication in the intensive care unit (ICU). It is difficult to diagnose in complex severely ill patients because specific biomarkers are lacking. The systemic inflammatory response syndrome (SIRS) is often associated with sepsis, but is also common after injury and surgery. We have developed an automatic algorithm for detection of SIRS-criteria (tachycardia, tachypnea, leukocytosis and fever). The change in algorithm-derived SIRS-descriptors was evaluated in sepsis cases and controls from a polytrauma cohort to describe and quantify the inflammatory process in both groups over time.

Methods: We identified a cohort of polytrauma patients from the electronic medical records of a tertiary care center's surgical ICU, 2006-2011. SIRS-criteria were determined with an electronic algorithm for each minute of the first 24 hours after admission and before sepsis therapy initiation, respectively. For each 60-minute interval, three SIRS-descriptors were calculated: (1) SIRS-criteria average for SIRS quantification over time, (2) first-to-last minute difference for trend detection, and (3) change count reflecting SIRS-criteria fluctuation. Time trends in SIRS-descriptors were statistically evaluated with linear mixed models for repeated measures with an autoregressive (ARMA (1,1)) correlation structure.

Results: Of 256 clinician validated polytrauma patients, 85 developed sepsis (33.2%). Excluding the first hour of data acquisition immediately after admission, SIRS-criteria average in both groups declined by 0.020 criteria on average per hour (95% CI 0.025-0.015). Change count of criteria decreased slightly in both groups (0.010 criteria per hour, 95% CI 0.013-0.006), indicating increasing stability. There was no change in either group in the criteria trend. In the 24-hour period before sepsis diagnosis the SIRS-criteria average increased by 0.026 criteria per hour (95% CI 0.017-0.035) in cases while it remained stable in controls. There was no meaningful hourly change in either trend or fluctuation in either group.

Conclusions: SIRS-descriptors, particularly SIRS-criteria average, are suggested as responsive measures for quantification of the systemic inflammatory process.

Simulating recurrent event data with a competing terminal event when failure rates depend on multiple time scales

Katharina Ingel, Fenja Nickel, Harald Binder, Antje Jahn-Eimermacher

Institute of Medical Biostatistics, Epidemiology and Informatics, University Medical Center of the Johannes Gutenberg-University Mainz

Simulation studies are an important tool for planning clinical trials with a complex time-to-event structure. For example in cardiovascular studies, patients are at an increased risk for disease-associated hospitalizations and for competing death. Motivated by clinical and cohort studies, we will present simulation algorithms for different multistate models for time to recurrent hospitalizations and competing death.

Simulation models usually define all transition intensities on one time scale. For the time scale being the time since the last event (gap time approach), the simulation is straightforward. For the time scale being the time since entry into the study or since beginning of the disease (total time approach) a recursive algorithm for nested competing risk models can be applied and has been implemented [1,2]. However, failure rates might also depend on both timescales. For example, for cardiovascular patients the risk for disease-associated hospitalizations will increase with age or disease duration (total time) but also treatment modifications as a response on hospitalization events will affect their further risk (gap time). Models to reflect these situations, e.g. the autoregressive model [3] and the trend-renewal process [4], will be extended to the nested competing risk model and corresponding simulation methods will be presented.

In most simulations studies either gap time or total time is considered as the underlying time scale. We propose methods to incorporate both time scales and illustrate these by an application in cardiovascular disease.

[1] Jahn-Eimermacher A, Ingel K, Ozga A, Preussler S and Binder H (2015): Simulating recurrent event data with hazard functions defined on a total time scale. BMC Medical Research Methodology 15:16.

[2] simrec: An R-package for Simulation of Recurrent Event Data.
https://github.com/katharinaingel/simrec

[3] Metcalfe C and Thompson SG (2006): The importance of varying the event generation process in simulation studies of statistical methods for recurrent events. Statistics in Medicine 25, 165-179.

[4] Pietzner D and Wienke A (2013): The trend-renewal process: a useful model for medical recurrence data. Statistics in Medicine 32, 142-152.

Authors: Jan Beyersmann, Christine Schrade

Title: Florence Nightingale, William Farr and competing risks

Abstract: Aiming for hospital sanitary reforms, Florence Nightingale (1820-1910), a pioneer of nursing and modern hospital hygiene, joined forces with William Farr (1807-1883), whose 1885 book Vital Statistics is an early important contribution to statistics for the life sciences. Today, Farr's book is well remembered for his clear recognition of time-dependent (immortal time) bias. He mockingly argued: "make [young men] Generals, Bishops, and Judges -for the sake of their health!" As in Vital Statistics, Nightingale and Farr used incidences densities to quantify hospital mortality, that is, the number of hospital deaths divided by the hospital population time at risk. Nightingale's and Farr's methodological choice was criticized by their contemporaries, one issue being that incidence densities may exceed 100%.

Interestingly, these questions are still with us today. In hospital epidemiology, quantifying incidence of in-hospital events typically uses incidence densities, although hospital mortality is nowadays almost exclusively expressed as the number of hospital deaths divided by the size of the hospital population, the so-called incidence proportion. In randomized clinical trials, the risk of adverse events is often quantified via incidence densities. However, their use is also criticized in this context, one issue being possibly informative censoring by the primary trial outcome. This talk outlines that the modern survival analytical concepts of competing risks and cause-specific hazards overcome such criticism and, e.g., reconcile the concepts of incidence densities and incidence proportion. We will show that Nightingale and Farr were essentially aware of these connections.

Authors:
Nadine Binder and Martin Schumacher
Institute for Medical Biometry and Statistics, Medical Center - University of Freiburg, Germany

Title:
Estimating hazard ratios in cohort data with missing disease information due to death

Abstract:
In epidemiological studies information on disease status can often only be collected at a few discrete follow-up times, often after some years. This can be done only retrospectively in individuals who are alive at follow-up, but the information will be missing for those who died between two visits. Right-censoring the death cases at the last visit (ad-hoc survival analysis) generally underestimates the disease incidence [1], resulting in biased hazard ratio estimates, but in both directions [2].
Since the observed data actually has underlying illness-death structure, we investigate to which extent three different multi-state approaches [1,3,4], taking the death cases into account, provide less biased hazard ratio estimates. We investigate to which extent the approaches allow to correct for the bias in simulation studies. We also evaluate them on a real data example, a random sample of the French elderly population-based PAQUID cohort study of mental and physical aging (study details in [1]), provided with the R package SmoothHazard [5], the latter which implements two of the three multi-state approaches.
All investigated approaches can reduce the bias to a different extent and under different circumstances.

References:
[1] Joly P, Commenges D, Helmer C, Letenneur L (2002). A penalized likelihood approach for an illness-death model with interval-censored data: application to age-specific incidence of dementia. Biostatistics 3(3):433-43.

[2] Binder N, Schumacher M (2014). Missing information caused by death leads to bias in
relative risk estimates. J Clin Epidemiol 67(10):1111-20.

[3] Leffondré K, Touraine C, Helmer C, Joly P (2013). Interval-censored time-to-event and competing risk with death: is the illness-death model more accurate than the Cox model? Int J Epidemiol 42(4):1177-86.

[4] Yu B, Saczynski JS, Launer L (2010). Multiple imputation for estimating the risk of developing dementia and its impact on survival. Biom J 52(5):616-27.

[5] Touraine C, Gerds TA, Joly P. The SmoothHazard package for R: Fitting regression models to interval-censored observations of illness-death models. Technical Report 12, University of Copenhagen, Department of Biostatistics 2013.

Differential mortality explains maximum of age-specific seroprevalence in Hepatitis A infections

Ralph Brinks, Sandra Landwehr, Annika Hoyer

Epidemiologic surveys of the age-specific prevalence of antibodies to hepatitis A virus (HAV) provide information on the spread of infection such as the infection rate (force of infection) and other age-dependent characteristics.

Common Markov chain models do not allow maxima in the age-specific seroprevalence and introduce diminishing immunity to HAV to explain the peak at the age of 50 observed in several countries (see [1] for a review and [2] for a more recent paper). There are several reasons against diminishing immunity at these early ages. For example, immunosenescence typically occurs at higher ages [3].

Using the methodology of Markov chains, we extend the illness-death model to the case of differential mortality between ever- and never-infected. Then, age-specific seroprevalence is the solution of a Riccati differential equation, whose solutions may have maxima.

Given that HAV infections are associated with low socioeconomic status, which in turn is associated with higher mortality, the new model with differential mortality explains the observed maxima without the concept of diminishing immunity.

The new model can be applied to other diseases, like dementia and diabetes [4], where differential mortality is common.

References

[1] Yang G, Chang MN (1990) A stochastic model for analysing prevalence surveys of Hepatitis A antibody, Mathematical Biosciences 98:157–69.

[2] Yang G (2013) Neyman, Markov processes and survival analysis, Lifetime Data Analysis, 19:393–411.

[3] Barcellini W, Borghi MO, Sguotti C (1988) Heterogeneity of immune responsiveness in healthy elderly subjects, Clin Immunol Immunopathol 47:142–51.

[4] Brinks R, Landwehr S (2015) Change rates and prevalence of a dichotomous variable: simulations and applications, PLOS One 10(3): e0118955.

Contact
Ralph Brinks
German Diabetes Centre
40225 Düsseldorf
ralph.brinks@ddz.uni-duesseldorf.de

Using pseudo-values for comparing long-term survival after stem-cell transplantation (SCT) with long-term survival after chemotherapy

Ulrike Pötschger[1,2], Harald Heinzl[2], Maria Grazia Valsecchi[3], Martina Mittlböck[2]

1 Children's Cancer Research Institute, Vienna, Austria

2 Center for Medical Statistics, Informatics, and Intelligent Systems, Medical University of Vienna, Vienna, Austria

3 Center of Biostatistics for Clinical Epidemiology, Department of Health Sciences, University of Milan-Bicocca, Monza, Italy

Aims
Allogeneic SCT is a therapeutic option in high-risk leukemia whose ability to improve long-term survival is still under discussion. The evaluation of this question is often based on donor availability, not known immediately when decision for SCT is made. From a methodological perspective, when investigating long-term survival rates, it is unclear and challenging how to adjust for waiting-time until donor identification without relying on proportional hazards.
Methods
The pseudo-value regression provides an attractive tool to model time-fixed covariates in presence of non-proportional hazards. A modification will be introduced to assess the effect of a binary time-dependent covariate like donor availability. Pseudo-values are generalized for patients with a suitable donor to estimate cumulative treatment effects using the cumulative hazards before and after transplantation. These generalized pseudo-values and common pseudo-values for patients without a suitable donor can be modelled in a weighted generalised linear model to assess treatment effects and include further baseline covariates in a straightforward manner. Real data in leukaemia and a simulation study illustrate the practical value and statistical properties of the novel approach.
Conclusion
The proposed novel approach allows investigating the impact of a binary time-dependent covariate on long-term survival without relying on proportional hazards. Furthermore, it allows to model and adjust for further baseline covariates.

Identifying relevant variables by simultaneous variable selection between dynamic prediction models based on pseudo-values

Daniela Zöller[1], Irene Schmidtmann[1], Arndt Weinmann[2], Thomas Gerds[3], Harald Binder[1]

1: University Medical Center Mainz, IMBEI, Mainz, Germany

2: University Medical Center Mainz, 1. Med, Mainz, Germany

3: University of Copenhagen, Department of Public Health, Copenhagen, Denmark

In a competing risks setting for time-to-event data the cumulative incidence is the absolute risk that the event of interest occurs within a certain time. It is an easily interpretable quantity and patients can be compared with respect to their cumulative incidences for different competing risks. By evaluating models for the cumulative incidence at a series of landmark time points, dynamic prediction models can be obtained. Although the idea of dynamic prediction is that some covariates are only relevant for some of the landmarks, there might be a set of covariates that should be included at all landmarks. One way to analyze effects on the cumulative incidence in right censored data is by using a pseudo-value approach, which allows, among others, the application of direct regression models. Using a suitable link function, the model allows for a straightforward interpretation of covariate effects on the cumulative incidence. We combine this approach with dynamic prediction to identify a potential set of covariates relevant for all landmark models and to deal with sparse information at early and late times of the follow-up.

Specifically we propose a stagewise regression algorithm. This algorithm provides estimates of possibly time-varying effects of covariates. The idea is to perform variable selection simultaneously across time points but allowing for separate estimates at each time point. The covariates relevant for all landmark models are identified with the help of inclusion frequencies. We apply the algorithm to clinical cancer registry data from hepatocellular carcinoma patients.

The combination of pseudo-values with dynamic prediction is seen to stabilize the variable selection between the time points and to improve interpretation of the results.

The average hazard ratio – a good effect measure for time-to-event endpoints when the proportional hazard assumption is violated?

Geraldine Rauch, Werner Brannath, Mathias Brückner, Meinhard Kieser

In many clinical trial applications, the endpoint of interest corresponds to a time-to-event endpoint. In this case, group differences are usually expressed by the hazard ratio which is the standard effect measure in a time-to-event setting. The commonly applied test is the logrank-test, which is optimal under the proportional hazard assumption. However, there are many situations in which the proportional hazard assumption is violated. Especially in applications were a full population and several subgroups or a composite time-to-first-event endpoint and several components are considered, the proportional hazard assumption usually does not simultaneously hold true for all test problems under investigation.

For the case of non-proportional hazards, Kalbfleisch and Prentice (1981) proposed the so-called average hazard ratio as an alternative effect measure which can also be used to construct a corresponding test for group differences. Furthermore the average hazard ratio has a meaningful interpretation even in the case of non-proportional hazards. However, it is hardly ever used in practice, whereas the standard hazard ratio is commonly reported in clinical trials analyzing time-to-event data.

The aim of this talk is to give a systematic comparison of the two approaches for different time-to-event settings corresponding to proportional and non-proportional hazards and illustrate the pros and cons of both methods in application. In particular, we will discuss how to apply the average hazard ratio in the presence of competing risks. Comparisons will be made by Monte-Carlo simulations and by a real clinical trial example.

Reference:

[1] Kalbfleisch D., Prentice R. L. Estimation of the average hazard ratio. *Biometrika* 1981, **68**:105-112.

Random forests for survival analysis using maximally selected rank statistics

Marvin N. Wright[1], Theresa Dankowski[1], Andreas Ziegler[1,2]

1 Institut für Medizinische Biometrie und Statistik, Universität zu Lübeck, Universitätsklinikum Schleswig-Holstein, Campus Lübeck, Germany

2 Zentrum für Klinische Studien, Universität zu Lübeck, Germany

We use maximally selected rank statistics to optimize the split points in random forests for survival analysis. The method minimizes p-values for association between split points and survival time to avoid split point selection bias. We describe maximally selected rank statistics for survival endpoints, several p-value approximations and the implementation of the proposed random forests approach. The different p-value approximation methods are compared in a simulation study. It is shown that unbiased split point selection is possible. However, there is a trade-off between unbiased split point selection and runtime. A benchmark study of prediction performance on six real and simulated datasets shows that the new method performs equally well or better than random survival forests, conditional inference forests and the Cox model. We conclude that random forests with maximally selected rank statistics are a useful alternative to other random forest approaches for survival analysis.