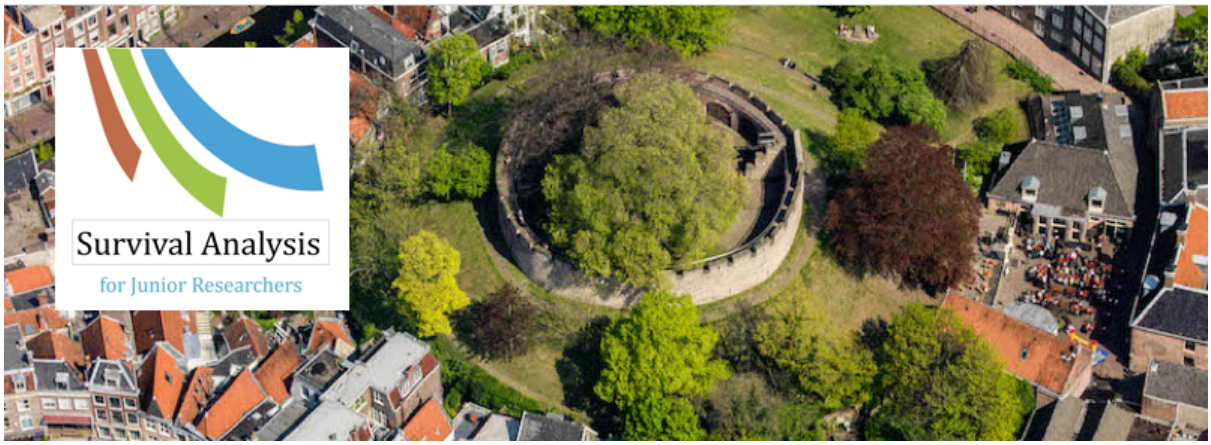


Survival Analysis for Junior Researchers Conference

Leiden, April 24th - 26th, 2018

Information and abstracts

Universiteit Leiden
The Netherlands

Leiden University Medical Center

Statistical Science

VVSOR

ROYAL STATISTICAL SOCIETY

WILEY

Contents

Welcome **1**

Conference Details **3**
- Map . . . 3
- Transportation . . . 4
- Conference Venue . . . 4
- Locations of Social Events . . . 5
- Special Issue of Statistica Neerlandica on papers presented at SAfJR2018 . . . 6
- Prizes for best presentation and poster . . . 7

Keynote Speakers **9**
- Prof. dr. Hein Putter (Leiden University Medical Center) . . . 9
- Dr. Maja Pohar Perme (University of Ljubljana) . . . 9
- Dr. Birgit Lissenberg-Witte (VU University Medical Center, Amsterdam) . . . 10

Sponsors **11**
- Prof. dr. Jacqueline Meulman, on behalf of the Statistical Science Master Programme 11
- Netherlands Society for Statistics and Operations Research . . . 11
- Royal Statistical Society . . . 11
- Wiley . . . 12

Organising Committee **13**
- Stéphanie van der Pas . . . 13
- Anja Rüten-Budde . . . 13
- Sanne Willems . . . 13

Poster List **15**

Abstracts **17**
- Keynote Speakers . . . 17
- Session 1: Causality . . . 19
- Session 2: Informative censoring . . . 23
- Session 3: Flexible parametric modelling . . . 26
- Session 4: Multiple time scales & multistate models . . . 29

Session 5: Pseudo-observations 35
Session 6: Dynamic prediction & competing risks 38
Poster session 44

Programme Details **65**

Welcome

Dear delegate,

Welcome to Leiden! We are delighted that the Survival Analysis for Junior Researchers conference has crossed the Channel for its 7th edition, hopefully starting a tradition where the conference is hosted in a new location every year. The conference brings together career-young statisticians working with time-to-event data and offers them a platform to meet their peers and to present their work. Many past delegates look back fondly to the times when they participated in the conference.

Thanks to the enthusiasm by many of you to contribute to the conference, we can again offer a program filled to the brim with talks, ranging from applied to theoretical, and covering many topics in survival analysis, such as informative censoring, dynamic prediction, and causality.

This year, Hein Putter, Maja Pohar Perme and Birgit Lissenberg-Witte have kindly agreed to give the keynote presentations. Besides the talks and posters, there will be many breaks, a welcome reception at the Leiden City Hall, and a conference dinner, to ensure that there will be plenty of opportunity for the delegates to interact with each other.

We hope that you will find the conference a valuable experience, and that you will enjoy your stay in The Netherlands!

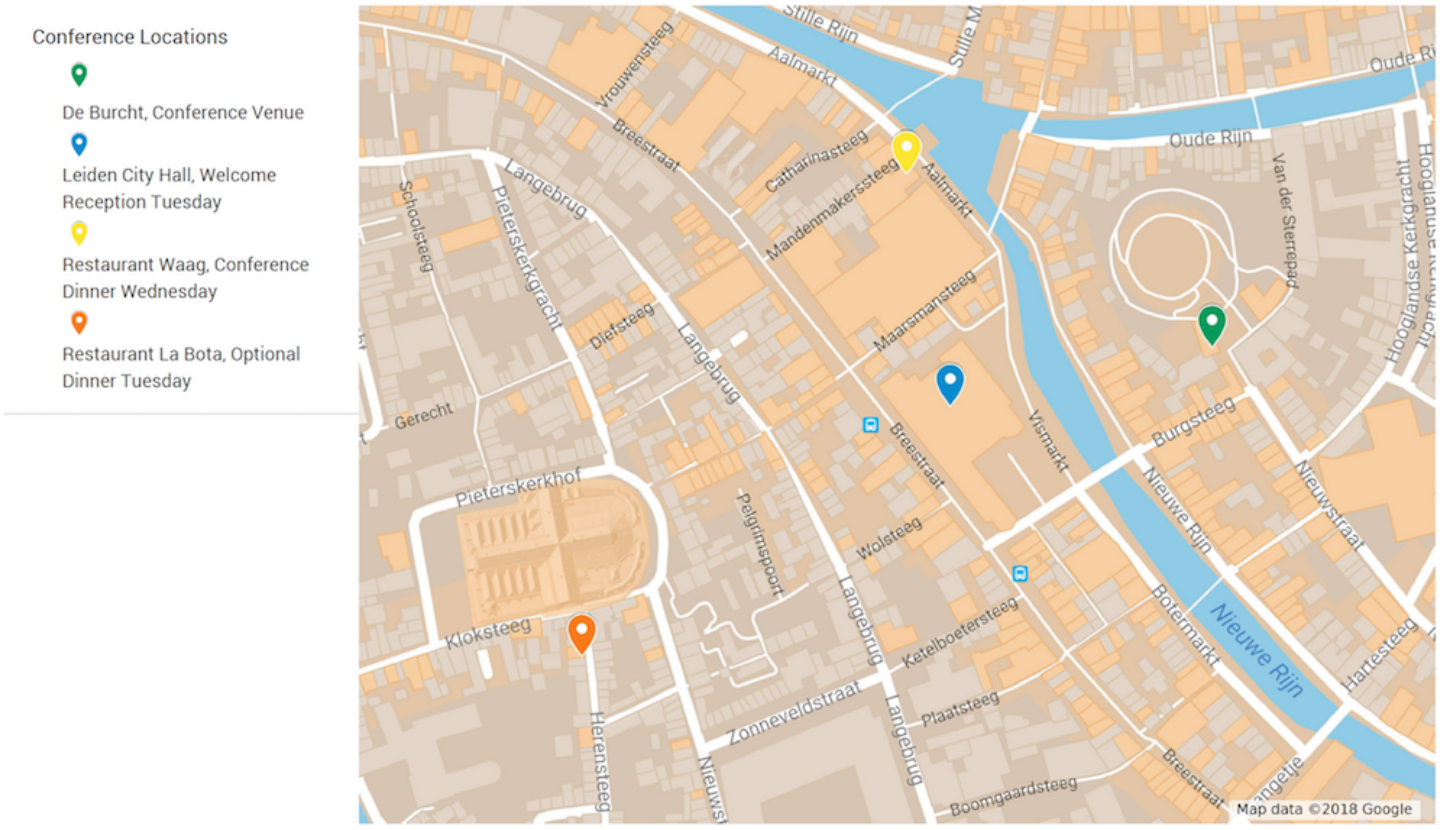
With best wishes and regards from the 2018 SAfJR Organising Committee,

Sanne, Anja and Stéphanie

E-mail us!
You can contact the organising committee via safjr2018@leidenuniv.nl.

Conference Details

Map



Map of Leiden

Transportation

Once in Leiden, you can easily move from place to place on foot; all conference locations are within a 10-minute radius. Please see the map on page 3.

Conference Venue

The conference venue is grand cafe De Burcht. This a restaurant next to one of Leiden's fort; one of the many charming parts of Leiden. All presentations and the poster session will be in the garden room. The entrance to this room is on the ground floor, next to the staircase that leads up to the grand cafe. Lunches will be served in the restaurant across the square.

Entrance to the garden room of the conference venue (below the arrow).

Locations of Social Events

Welcome reception at Leiden city hall (Tuesday April 24th, 2018)

The mayor of Leiden will welcome all SAfJR2018 participants during a reception at the city hall, just a 2-3-minute walk from the conference venue.

Address: Stadhuisplein 1 (see map on page 3)

Optional dinner at restaurant La Bota (Tuesday April 24th, 2018)

There is an optional dinner after the welcome reception. The costs for this dinner are €20,- and include a 3-course meal and 2 drinks at restaurant La Bota. Until April 16th, participants can register for this dinner via the online form sent to them via e-mail and pay for this dinner at the registration desk. Only cash payments are accepted.

Address: Herensteeg 9 (see map on page 3)

Conference dinner at restaurant Waag (Wednesday April 25th, 2018)

The conference dinner will be at restaurant Waag, a 4-5 minute walk from the conference venue. Our group will have dinner in a separate dining room, which you can access via the side entrance in the "Mandenmakerssteeg" (the alley on the right of the main entrance of the restaurant).

Address: Aalmarkt 21 (see map page 3)

Special Issue of Statistica Neerlandica on papers presented at SAfJR2018

A selection of papers presented at Survival Analysis for Junior Researchers will be featured in a Special Issue of Statistica Neerlandica.

The editors of Statistica Neerlandica, the Journal of the Netherlands Society for Statistics and Operations, published by Wiley, have expressed their interest in publishing a Special Issue on topics presented at SAfJR2018. After the conference, a selection of presenters will be personally invited to contribute a full paper based on their conference presentation for this Special Issue. The deadline for submission to this Special Issue will be August 1st. Every contributor will be invited to review one of the other contributions. More information on the journal can be found on the website of the Netherlands Society for Statistics and Operations, and on the publisher's website.

https://www.vvsor.nl/publications/statistica-neerlandica/

https://onlinelibrary.wiley.com/journal/14679574

Prizes for best oral presentation and poster

Netherlands Society for Statistics and Operations Research and Wiley have generously provided the prizes for the best poster and best oral presentation. The prize winner for the best poster will be awarded a Wiley giftcard of £100,- and the best presenter will be awarded an Amazon giftcard of €100,-.

Keynote Speakers

Prof. dr. Hein Putter (Leiden University Medical Center)

Prof. dr. Hein Putter is professor in the Biomedical Data Sciences department at the LUMC. He collaborates closely with the Department of Surgery and the Department of Oncology of the LUMC, and with international organizations like the European Organisation for the Research and Treatment of Cancer (EORTC) and the European Group for Blood and Marrow Transplantation (EBMT). His most important research area is survival analysis. Within this field, his main research interests are competing risks and multi-state models, dynamic prediction, and frailties. The topics competing risks/multi-state models and dynamic prediction are closely related. Together with Hans van Houwelingen, he wrote the book "Dynamic Prediction in Clinical Survival Analysis". He also co-wrote the mstate-package in R for estimation and prediction in non- and semi-parametric multi-state and competing risks models.

Prof. dr. Hein Putter will give us a tutorial on multi-state models in survival analysis.

Dr. Maja Pohar Perme (University of Ljubljana)

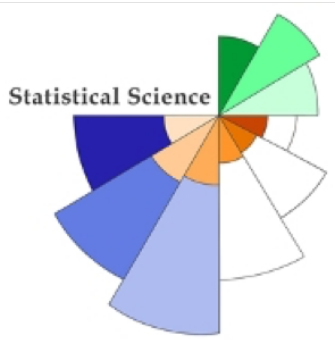
Dr. Maja Pohar Perme is an associate professor of Biostatistics at the Medical faculty of the University of Ljubljana and teaches both medical students as well as statistics students at masters and PhD level. As statistician, she is the co-editor of the Slovenian Medical Journal. Her research focuses on survival analysis with particular interest in relative survival methodology. She is researching properties of the methods traditionally used in the relative survival setting and has made several contributions to the field by introducing methodology for non-parametric estimation, semi-parametric modelling, goodness of fit testing, etc. Her current research focus is the use of pseudo-observations in relative survival. She is the author of the relsurv-package in R.

Dr. Birgit Lissenberg-Witte (VU University Medical Center, Amsterdam)

Dr. Birgit Lissenberg-Witte aims her current research at the VU University Medical Center at more applied methodology in health science. She developed algorithms for more accurate estimation of cumulative incidence by correcting for misclassification of events due to limited accuracy of screening tests. Furthermore, she is involved in clinical decision research as a statistical consultant. Birgit teaches in the area of statistical methodology. During her PhD research at the Delft University of Technology, she developed nonparametric methods to estimate the cumulative incidence of asymptomatic diseases, in particular the incidence of viral infections (HIV and Hepatitis A), and focused on mathematical properties of these estimation methods.

Sponsors

This conference is sponsored by:



Prof. dr. Jacqueline Meulman, on behalf of the Statistical Science Master Programme

The master Statistical Science provides students with a thorough introduction to the general philosophy and methodology of statistical modelling and data analysis and focusses on applications in the Life and Behavioural Sciences or Data Science.

Netherlands Society for Statistics and Operations Research

The Netherlands Society for Statistics and Operations Research (VVSOR) is the Dutch society for statistics and operations research. This society was founded on August 15th 1945 and aims to stimulate the application and study of statistics and operations research, in favour of science and society.

Royal Statistical Society

The Royal Statistical Society (RSS) is one of the world's leading organisations to promote the importance of statistics and data.

Wiley

Wiley is a global publishing company that has been helping people and organisations develop the skills and knowledge they need to succeed for over 200 years.

Organising Committee

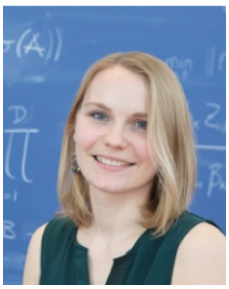
This year's SAfJR conference is organised by three statisticians working at Leiden University and Leiden University Medical Center.

Stéphanie van der Pas

Stéphanie van der Pas is an assistant professor in statistics at the Mathematical Institute of Leiden University and at the Leiden University Medical Centre, and a guest fellow at the Princess Máxima Center for Pediatric Oncology. Stéphanie's research interests include survival analysis, Bayesian nonparametrics, high-dimensional sparse data, and translating theoretical statistical results to medical practice.

Anja Rüten-Budde

Anja is a PhD candidate at the Mathematical Institute of Leiden University. Her research addresses predictive performance indicators for survival models. Motivation for this research are the applications of prediction models in the medical field, particularly in oncological data. This research has been funded by the KWF.



Sanne JW Willems

Sanne is a PhD candidate in Applied Statistics at the Mathematical Institute of Leiden University. Her PhD research focusses on the analysis ordinal variables in models fitted with a maximum likelihood approach. Furthermore, she is interested in science communication and plans to do a research project in this field as part of her PhD project.

Poster List

	Posters
1.	Lizbeth Burgos Ochoa: *Comparison of methods to integrate mediation and survival analysis*
2.	Yuntao Chen: *A varying-coefficient multi-state model to evaluate the effect of a longitudinal biomarker on a disease process*
3.	Elinor Curnow: *Multiple imputation strategies for interval-censored event times*
4.	Carla Díaz-Louzao: *Flexible Bayesian Additive Joint Models for Longitudinal and Time-to-Event Data: Applicacion to Liver Transplantation Data*
5.	Oswaldo Gressani: *P-splines and Laplace approximations for fast Bayesian inference in a flexible promotion time cure model*
6.	Sarwar Islam Mozumder: *Restricted mean lifetime estimation in the presence of competing risks from within the flexible parametric modelling framework*
7.	Kévin Jaunâtre: *Extreme value estimation with Cox model*
8.	Radoslav Kovar: *Dependent competing risks in a mixture cure model*
9.	Carlo Lancia: *Target, Anticipated, and Regulated Received Dose Intensity*
10.	Romin Pajouheshnia: *A comparison of methods to account for time-varying treatment use when developing a prognostic model*
11.	Eunyoung Park: *Penalized Variable Selection in Accelerated Failure Time Models with Random Effects*
12.	Douglas Stram: *Novel statistical methods for efficient identification of biomarkers for personalized cancer treatment*
13.	Leili Tapak: *Sparse Variable selection for competing risks*
14.	Lucy Teece: *Competing Risks in Pregnancy Prognosis: A study on the effect of delivery as a competing event for antenatal adverse events in women with early-onset pre-eclampsia*

15. Guillermo Villacampa: *Multi-state models to evaluate the benefit of a surgical procedure*

16. Maja von Cube: *Estimation of the population-attributable fraction for survival settings with time-dependent exposures*

17. Susanne Weber: *Accounting for Length of Stay in Regression Models in Hospital Epidemiology*

Abstracts

Keynote Speakers

Simplifying survival analysis with pseudo-observations

Dr. Maja Pohar Perme[1]

[1]Institute for Biostatistics and Medical Informatics, Faculty of Medicine, University of Ljubljana, Slovenia

Abstract:
Pseudo-observations present a general tool that can simplify a wide range of analyses in the survival field. The basic idea is to first handle the issue of censoring by defining an outcome that is available for each individual at each follow-up time. Having defined pseudo-observations, the basic need for special survival analysis methods is removed and one can use standard approaches available outside the survival field. We shall consider three different problems of survival analysis where pseudo-observations provide simple and flexible solutions: graphical methods for assessing goodness-of-fit for regression models, estimation in relative survival and regression with years lost or saved as the outcome of interest. We will use these examples to discuss the properties of pseudo-observations, their advantages and their issues. While it shall often turn out that compared to model specific solutions, pseudo-observation approach may be less efficient, their importance becomes clear when considering models for which no method has yet been developed. We shall use both simulated and real data to illustrate each of the examples.

Nonparametric estimation of the cumulative incidence function: accounting for imperfect tests in interval censored data

Dr. Birgit I. Lissenberg-Witte[1]

[1]Department of Epidemiology and Biostatistics, VU University Medical Center Amsterdam, Amsterdam Public Health Research Institute

Abstract:
In many longitudinal clinical studies, the primary outcome measure is the time until a certain event occurs. An example of such an event is a recurrent lesion after surgery. If a patient is tested for a recurrent event at several follow-up time points, the time-to-event is subject to interval censoring. Depending on the number of follow-ups, data can be modelled by the case K interval censoring model (each patient is tested exactly K times) or the mixed case interval censoring model (each patient can have a different number of tests, with a pre-defined maximum). Since only the times of the last negative and first positive test are sufficient for estimating the survival function (or equivalently the cumulative incidence function), in both models the data can be reduced to case 2 interval censored data where patients are only tested twice.
A natural estimator for the cumulative incidence function of the time-to-event is the non-parametric maximum likelihood estimator (NPMLE) for the case 2 interval censoring model. However, in many clinical situations, the occurrence of the event can only be assessed by a non-invasive diagnostic test with a limited sensitivity and/or specificity. The cumulative distribution function (or equivalently the survival function) will be under- or overestimated if the estimation method is not corrected for misclassification of the events.
First, I will describe the mixed case interval censoring model in more detail and describe several algorithms to estimate the cumulative distribution function. Then, I will incorporate the possible imperfect diagnostic test in the model, and describe an expectation-maximization (EM)-algorithm to compute the NPMLE mixed case interval censoring model. The algorithm has a closed-form solution for the combined expectation and maximization step and is computationally undemanding. In a simulation study, I compare the NPMLE to the uncorrected NPMLE that ignores possible misclassification. Finally, I illustrate the algorithm with follow-up data from women treated for cervical precancer.

Session 1: Causality

Chair: Cristina Boschini

A simple to implement approach for estimating attributable in-hospital mortality of a time-varying exposure

Johan Steen[1], Stijn Vansteelandt[2,3], Johan Decruyenaere[1]

[1]Department of Intensive Care, Universitair Ziekenhuis Gent, Belgium
[2]Department of Applied Mathematics, Computer Science and Statistics, Universiteit Gent, Belgium
[3]Department of Medical Statistics, London School of Hygiene and Tropical Medicine, UK

E-mail for correspondence: johan.steen@uzgent.be

Abstract:
Hospital-acquired infections form a major public health problem in developed countries. However, assessing their causal impact remains a subtle issue, especially from observational data, because patients susceptible to those infections generally experience poorer health conditions. Estimation of attributable in-hospital mortality of these infections thus necessitates confounding adjustment for severity of illness indicators. It is often less well understood that adjustment should not only be made at baseline, but also for the evolution of such indicators during hospitalization in order to avoid immortal time bias that might arise due to failure to account for the time-dependent nature of acquiring infection (Suissa, 2003). Assessment of the attributable fraction over time requires comparison of the observed cumulative incidence function with its non-observable counterfactual counterpart that would have been observed if all considered patients would either have died or been discharged without infection. Routinely collected health data that contain daily measurements of relevant time-varying confounders increasingly allow to meet reasonable assumptions under which bias-free estimation of the latter can be obtained. We propose an inverse probability weighted version of the extended Aalen-Johansen estimator (Westreich et al., 2010) for this counterfactual quantity and illustrate that existing R packages enable a simple to implement estimation approach.

Key words: attributable fraction; inverse probability weighting; competing risks; time-varying exposure; electronic health records.

S. Suissa (2003). Effectiveness of inhaled corticosteroids in chronic obstructive pulmonary disease: Immortal time bias in observational studies. American Journal of Respiratory and Critical Care Medicine, 168(1):49–53.

D. Westreich et al. (2010). Time scale and adjusted survival curves for marginal structural Cox models. American Journal of Epidemiology, 171(6):691–700.

Estimating the impact on survival of immediate vs delayed kidney transplantation

C. Olarte [1,2], E. Goetghebeur [1]

[1]Department of Applied Mathematics, Computer Science and Statistics, Faculty of Science, Ghent University, Ghent, Belgium.
[2]Department of Clinical Epidemiology, Biostatistics, Bioinformatics, Academic Medical Center, Amsterdam, The Netherlands.

E-mail for correspondence: Camila.OlarteParra@UGent.be

Abstract:
The impact on survival of immediate versus delayed kidney transplantation (preceded by dialysis), has been estimated from kidney transplant registries which follow patients from transplantation onwards. The treatment strategy is however decided when end-stage kidney disease is diagnosed. Naturally, the causal effect of the choice of treatment on survival, compares survival times under both strategies from this point onwards in exchangeable patient groups. This requires adjustment for baseline confounders at that same time and not for covariate values at the time of delayed transplantation which may be intermediate on the causal pathway from treatment choice to outcome. In this talk, we examine how structural accelerated failure models and G-estimation allow to estimate the survival benefit of kidney transplantation after varying lengths of dialysis, with data that follow patients from diagnosis. In addition, we consider the marginal effect of the choice of treatment on survival beyond end-stage kidney disease and explore the current magnitude of bias.

Key words: Structural Accelerated Failure Models; G-estimation; Survival; Time-zero; Kidney Transplantation.

D. Abramowicz, M. Hazzan, U. Maggiore, L. Peruzzi, P. Cochat, R. Oberbauer et al (2013). Does pre-emptive transplantation versus post start of dialysis transplantation with a kidney from a living donor improve outcomes after transplantation? A systematic literature review and position statement by the Descartes Working Group and ERBP. Nephrology Dialysis Transplantation, 31(5):691–7.

MA. Hernan, BC. Sauer, S. Hernandez-Diaz, R. Platte, I. Shrierg (2016). Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. Journal of Clinical Epidemiology, 79:70–75

DM. Vock, AA. Tsiatis, M. Davidian, EB. Laber, WM. Tsuang, CAF. Copeland et al (2013). Assessing the Causal Effect of Organ Transplantation on the Distribution of Residual Lifetime. Biometrics, 69: 820–829.

Sibling models in the Cox framework: A panacea to inferring causality from observational data or a new source of problems?

Anne Helby Petersen[1]

[1]Section of Biostatistics, Department of Public Health, University of Copenhagen, Copenhagen, Denmark

E-mail for correspondence: ahpe@sund.ku.dk

Abstract:
In epidemiology, sibling-stratified models have become an increasingly popular tool for reaching causal conclusions from observational data. These models circumvent the typical limitations of a non- randomized data collection by addressing intra-family differences, thereby controlling for unobserved, familial confounding by design. In the case where most unobserved confounding is shared within a family, this can allow for estimating causal effects. However, the combination of the sibling-design and time-to-event modeling comes with a few caveats. These include reductions in statistical power due to the reliance on intra-family variation, a dramatic increase in the computational burden and a limited transparency as to when the sibling design can appropriately be applied without introducing new - and bigger - obstacles into the search for unbiased causal effect estimates. In this presentation, the theoretical and practical consequences of these caveats are discussed with reference to an application of two sibling survival models - the stratified Cox model and the novel between-within gamma frailty Cox model (Sjölander et al. 2013) - in a register-based cohort study investigating the causal effects of mode of delivery by cesarean section on the risk of subsequent development of psychiatric disorders in the offspring (Axelsson et al. 2017).

Key words: Sibling designs; causal inference; observational data; Cox regression; between-within model.

Paul Axelsson, Tine Clausen, Anne H. Petersen and 7 additional authors (2017). Neither delivery by cesarean section nor antibiotic use in early infancy increases the risk of ADHD later in life? A national cohort study using the between-within survival model for siblings. Submitted for JAMA.

Arvid Sjölander, Paul Lichtenstein, Henrik Larsson and Yudi Pawitan (2013). Between-within models for survival analysis. Statistics in Medicine 32(18):3067-3076.

Session 2: Informative censoring

Chair: Regina Stegherr

Modeling recurrent event times subject to right-censoring with D-vine copulas

Nicole Barthel[1], Candida Geerdens[2], Claudia Czado[1], Paul Janssen[2]

[1] Department of Mathematics, Technische Universität München, Garching, Germany
[2] Center for Statistics, Universiteit Hasselt, Hasselt, Belgium

E-mail for correspondence: nicole.barthel@tum.de

Abstract:
In several time-to-event studies, the event of interest is recurrent and the data for a sample unit corresponds to a series of gap times between subsequent events. A limited follow-up period causes the last gap time to be subject to right-censoring. Due to the sequential nature of the data dependence is induced such that gap times and censoring times cannot be assumed independent. Also, the number of occurrences may differ between sample units, making gap time data unbalanced. To unravel the dependence pattern of gap times, these data features need to be taken into account. D-vine copulas serve this purpose and naturally capture the serial dependence inherent in gap time data. One- and two-stage likelihood based estimation is discussed. In the two-stage procedures nonparametric marginal estimation under dependent right-censoring is introduced. Hence, existing work is extended in several directions: for recurrent time-to-event data, so far only parametric survival margins and the restrictive class of Archimedean copulas have been considered. Simulations are used to evaluate the finite sample performance of the proposed estimation strategies. The analysis of real data on recurrent asthma attacks in children provides new and relevant insights by showing how the strength and type of dependence changes over time.

Key words: Dependence modeling; D-vine copulas; Induced dependent right-censoring; Maximum likelihood estimation; Unbalanced recurrent event time data.

Barthel, N., Geerdens, C., Czado, C. and Janssen, P. (2018). Modeling recurrent event times subject to right-censoring with D-vine copulas. working paper.

Barthel, N., Geerdens, C., Killiches, M., Janssen, P. and Czado, C. (2017). Vine copula based likelihood estimation of dependence patterns in multivariate event time data. Computational Statistics and Data Analysis, doi:10.1016/j.csda.2017.07.010.

Joint models for survival and longitudinal data when the observation process is informative

A Gasparini[1], J Barrett[2], KR Abrams[1], MJ Crowther[1]

[1]Biostatistics Research Group, Department of Health Sciences, University of Leicester, Leicester, United Kingdom
[2]MRC Biostatistics Unit, University of Cambridge, Cambridge, United Kingdom

E-mail for correspondence: ag475@leicester.ac.uk

Abstract:
Electronic health records are being increasingly used in medical research to answer more relevant and detailed clinical questions; however, they pose new and significant methodological challenges. For instance, observation times are likely correlated with the underlying disease severity: patients with worse conditions utilise healthcare more and have worse biomarkers recorded. Additionally, a terminal event truncating observation of the longitudinal process is deemed informative when it correlates with disease severity. Traditional methods for analysing longitudinal data assume independence between observation times and disease severity; yet, with healthcare data such assumption unlikely holds leading to biased model estimates. Joint models for longitudinal and survival data can account for informative dropout processes, but research is scarce on whether inference is valid when the observation process is informative. Through extensive simulation studies, we compare different analytical approaches proposed to account for an informative visiting process. We cover both simple (including the number of measurements as a covariate) and complex methods (e.g. trivariate joint models for the longitudinal, survival, and visiting processes; Liu, 2008). We conclude by summarising which methods lead to valid inference, and under which settings, and describe how to fit the more complex models within an extended joint modelling framework (Crowther, 2017).

Key words: Healthcare data; Big data; Informative observation process; Joint models for longitudinal and survival data; Extended multivariate generalised mixed effects models.

L Liu, X Huang, and J O'Quigley (2008). Analysis of longitudinal data in the presence of informative observational times and a dependent terminal event, with application to medical cost data. Biometrics 64:950-958

MJ Crowther (2017). Extended multivariate generalised linear and non-linear mixed effects models. arXiv:1710.02223

Drop-outs in studies with a recurrent event endpoint: Consequences of violating the independent censoring assumption

Gerrit Toenges[1], Florian Voss[2], Antje Jahn-Eimermacher[1,3]

[1]Institute for Medical Biostatistics, Epidemiology and Informatics (IMBEI), University Medical Center Mainz, Mainz, Germany
[2]Boehringer Ingelheim Pharma GmbH & Co. KG, Ingelheim, Germany
[3]Darmstadt University of Applied Sciences, Darmstadt, Germany

E-mail for correspondence: gtoenges@uni-mainz.de

Abstract:
This work is motivated by clinical trials in chronic obstructive pulmonary disease (COPD), where treatment-efficacy is assessed by the analysis of recurrent exacerbations. The patient's rate of disease exacerbations may influence his decision on treatment discontinuation. If the exacerbation rate does not only depend on known covariates, but also on unmodeled patient characteristics, the independent censoring assumption that classical models for recurrent events rely on might be violated. Recently, a joint frailty proportional hazards model has been proposed to overcome that issue (Rogers et al., 2016).
Fitting a joint frailty model requires input parameters, which are crucial for convergence of the estimates but might be difficult to pre-specify. We therefore identified those parameters that contribute to the difference between estimates derived from joint frailty and classical models. We investigate the asymptotic bias of classical model estimates in the situation where the drop-outs are associated with the recurrent events according to a joint frailty model.
We show that direction and degree of differences in effect estimates can be analytically derived. In particular, we also identify situations where hazard ratio estimates of classical models are still unbiased despite model misspecification. These results are related to findings of a COPD case study.

Key words: Recurrent events; Joint frailty model; Informative Drop-out's; Marginal model; Bias

Rogers JK, Yaroshinsky A, Pocock SJ, Stokar D and Pogoda J (2016). Analysis of recurrent events with an associated informative dropout time: Application of the joint frailty model. Statistics in Medicine , 35:2195-2205.

Session 3: Flexible parametric modelling

Chair: Stuart Lacy

Building and Validating Flexible Parametric Survival Models

Sarah Booth[1], Mark J. Rutherford[1], Paul C. Lambert[1,2]

[1]Biostatistics Research Group, Department of Health Sciences, University of Leicester, Leicester, UK
[2]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

E-mail for correspondence: sb824@le.ac.uk

Abstract:
Building a model using a large dataset often results in many of the covariates being statistically significant. This can lead to selecting a highly complex model that is not clinically useful. Therefore, it is of interest to investigate the effect that adding a new covariate has on the predicted survival probabilities. Different methods for assessing the impact of a new covariate on the outcome will be discussed.
Previous research has found that survival models are rarely validated. However, in order for the models to be clinically useful they should be externally validated to verify that the model predicts survival well using other datasets, for example, from a different geographical area or time period.
The advantages and disadvantages of various techniques for performing external validation will be discussed and demonstrated using flexible parametric models fitted to population-based cancer data from England. Methods for recalibrating the model, such as adjusting the baseline hazard rate for different regions of England, will also be discussed.

Key words: Flexible parametric models; External validation; Discrimination; Calibration.

R. D. Riley, J. Ensor, K. I. E. Snell et al. (2016). External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges. BMJ, 353:i3140.

P. Royston and W. Sauerbrei (2004). A new measure of prognostic separation in survival data. Statistics in Medicine, 23(5):723-748.

E. W. Steyerberg, K. G. M. Moons, D. A. van der Windt et al. (2013). Prognosis Research Strategy (PROGRESS) 3: Prognostic Model Research. PLoS Med 10(2):e1001381.

Estimating age-at-onset distribution of the asymptomatic stage of a disease

P. Vart, M.A. Jonker

Department for Health Evidence, section Biostatistics, Radboudumc Nijmegen, the Netherlands

E-mail for correspondence: Marianne.Jonker@radboudumc.nl

Abstract:
Information on the age-at-onset of the asymptomatic stage of a disease can be of paramount importance in early detection and timely management of that disease. However, accurately estimating the distribution of the age-at-onset of the asymptomatic stage is hard, because the asymptomatic stage is difficult to detect for the patient and is often detected as an incidental finding or in case of recommended screening. In this presentation, we consider the estimation of age-at-onset of the asymptomatic stage of a genetic disease for carriers of the genetic causal variant. The carriers are identified through genetic screening of the family of symptomatic patients with the variant and their disease status was determined by a medical examination. Since the exact age-at-onset of asymptomatic stage is never observed, it is interval censored. By maximizing the conditional likelihood (conditional on the ascertainment event), an asymptotically unbiased maximum likelihood estimator can be obtained for the age-at-onset of the asymptomatic stage. The methodology is applied to data on a muscle disease. Data was fitted with a parametric Cox-regression model with a Weibull baseline hazard function that includes characteristics of the genetic variant as a covariate.

Key words: Interval censored; Pedigree data; Age-at-onset distribution; Conditional maximum likelihood.

Estimating cumulative incidences using flexible parametric models at individual and population levels

D.K. Kipourou[1], B. Rachet[1], H. Charvat[2], A. Belot[1]

[1]Cancer Research UK Cancer Survival Group, Faculty of Epidemiology and Population Health, Department of Non-Communicable Disease Epidemiology, London School of Hygiene & Tropical Medicine, Keppel St, London WC1E 7HT, UK
[2]Division of Prevention, Center for Public Health Sciences, National Cancer Center, Tokyo, Japan

E-mail for correspondence: dimitra-kleio.kipourou@lshtm.ac.uk

Abstract:
In competing risks, we account for death according to a specific cause and the quantities of interest are usually the cause-specific hazards and the cumulative incidence functions (CIFs). A given CIF could be obtained with a combination of the cause-specific hazards or via the subdistribution hazard. Here, we modelled the cause-specific hazards with flexible parametric models (FPM) using B-splines for the log baseline hazard and the time-varying effects, as implemented in the R-package mexhaz. We derived the variance of the CIFs at the population level using the multivariate delta-method. We conducted a simulation study to evaluate the performance of this approach in its ability to estimate the CIF using different functions for the cause-specific log baseline hazard and with or without time-dependent (TD) effect. In the scenario with TD effect, we tested both well- and miss-specified models. We showed that the FPM perform nearly as well as the non-parametric method, if we allow enough flexibility for the baseline hazards. Moreover, neglecting the TD hardly affects the CIF estimates of the whole population, but impacts them in the various subgroups.

Key words: competing risks; flexible parametric models; cause-specific hazard; cumulative incidence function

Session 4: Multiple time scales & multistate models

Chair: Camille Sabathé

A Non-standard Case-Control Design With A View Towards Rare Time-dependent Exposures in Hospital Epidemiology

J. Feifel[1], J. Beyersmann on behalf of COMBACTE-CARE consortium[1]

[1]Institute of Statistics, Ulm University, Helmholtzstrasse 20, 89081 Ulm, Germany

E-mail for correspondence: jan.feifel@uni-ulm.de

Abstract:
For large cohort studies with rare outcomes the nested case-control design is favorable due to an efficient use of limited recourses. Small subsets of the individuals at risk are randomly sampled at the observed event times and a weighted, stratified analysis takes over the role of the full cohort analysis. We propose a non-standard nested case-control design, where the sampling probability evolves over time and may even depend on the cohort and sampling history. This is especially helpful when the outcome is not necessarily rare, but interest lies in the impact of a time-dependent exposure such as the occurrence of an adverse event or disease progression. The martingale arguments underlying both the standard and non-standard nested case-control design allow for choosing controls in a more sophisticated manner than simple random sampling which is usually applied in practice. We will discuss several options how to account for past time-dependent exposure status within a nested case-control design and their relative merits. It will be seen that a smart utilization of the available information at each point in time can lead to a more powerful and simultaneously less expensive design. The methods will be illustrated by observational data on the impact of hospital-acquired infection on hospital mortality.

Key words: Cox proportional hazards model; Illness-death model; Nested case-control; Time-dependent inclusion weights.

Modelling the impact of time to the intermediate event in the illness-death model

Davide Paolo Bernasconi[1], Elena Tassistro[1], Maria Grazia Valsecchi[1], Laura Antolini[1]

[1]School of Medicine and Surgery, University of Milano-Bicocca, Monza, Italy

E-mail for correspondence: davide.bernasconi@unimib.it

Abstract:
The illness-death model is the simplest multistate model where the transition from the initial state to the final state (death) may be direct or involve an intermediate state (illness). The analysis of the impact of the transition to illness and that of the time to this transition on the hazard of death has a key role in gaining insights into the dynamic of the disease. The standard approach is to jointly model the hazards of illness and death measuring time on the clock-forward scale and including as covariates the time-varying illness status and either the time to illness (waiting time) or the time since illness (sojourn time). Another recently proposed approach based on the clock-forward scale simultaneously includes as covariates both the waiting time and the sojourn time. We propose a further approach based on a double time scale: time is measured on the clock-forward scale before transition to illness and on the clock-reset scale after this transition and only the waiting time is included as a covariate. We compare these models through a simulation protocol, showing that the double-scale model is the most appropriate to deal with semi-Markov and extended semi-Markov scenarios.

Key words: Illness-death model; Poisson model; time scales; transition hazard; waiting time.

P.K. Andersen, S. Esbjerg, T.I. Sorensen (2013). Multi-state models for bleeding episodes and mortality in liver cirrhosis. Statistics in Medicine, 19(4):587-599.

S. Iacobelli and B. Carstensen (2013). Multiple time scales in multi-state models. Statistics in Medicine, 32:5315-5327.

C. Eulenburg, S. Mahner, L. Woelber and K. Wegscheider (2015). A systematic model specification procedure for an illness-death model without recovery. PLoS ONE, 10(4):e0123489.

Time-dependent ROC Curve Estimation with Cure Fraction

K.M. Beyene[1], A. El Ghouch[1], A. Oulhaj[2]

[1]Institute of Statistics, Biostatistics and Actuarial Sciences, Catholic University of Louvain, Belgium.
[2] Institute of Public Health, United Arab Emirates University, United Arab Emirates.

E-mail for correspondence: kasu.beyene@uclouvain.be

Abstract:
The ROC curve and its summary measure AUC are the two commonly used tools to evaluate the classification accuracy of a continuous variable for a binary outcome. The time-dependent ROC curves have been used to assess the predictive ability of diagnostic markers for survival analysis. Several authors have proposed methods to estimate the time-dependent ROC curves and AUC for a survival analysis. However, the validity of the estimators of these methods relies on some assumptions. One of the assumptions is that, all subjects of the study population are susceptible to the event of interest and will eventually experience this event if the follow-up period is sufficiently long. However, this assumption may not be valid in many cases and hence studying the sensitivity of the estimators for the violation of this assumption is of substantial interest. The main aim of this article is to assess the validity of the time-dependent ROC curves estimator for data with cured subjects. To study the performance of the estimator, we conducted in-depth simulations. The simulation studies make evident that, when the marker is known or correctly estimated, the simple method of Li et al. (2016) is insensitive to the violation of the above assumption.

Key words: Mixture cure model; Simple method; time-dependent AUC; Promotion time model.

L. Li, B. Hu and T. Greene (2016). A simple method to estimate the time-dependent ROC curve under right censoring. Statistical Methods in Medical Research, 0(0):1-15.

M.S. Pepe (2003). The statistical evaluation of medical tests for classification and prediction. Oxford University Press.

M. Schmid and S. Potapov (2012). A comparison of estimators to evaluate the discriminatory power of time-to-event models. Statistics in Medicine, 31(23):2588-2609.

M. Schmid, H.A. Kestler and S. Potapov (2013). On the validity of time-dependent AUC estimators. Briefings in Bioinformatics, 16 (1):153-168.

N.A. Kamarudin, T. Cox and R. Kolamunnage-Dona (2017). Time-dependent ROC curve analysis in medical research: current methods and applications. BMC Medical Research Methodology, 17. ISSN:1471-2288.

P. J. Heagerty and Y. Zheng (2005). Survival model predictive accuracy and ROC curves. Biometrics, 61(1):92-105.

P.J. Heagerty, T. Lumley and M.S. Pepe (2000). Time-dependent ROC curves for censored survival data and a diagnostic marker. Biometrics, 56(2):337-344.

P. Blanche, J.F. Dartigues and H. Jacqmin-Gadda (2013). Review and comparison of ROC curve estimators for a time-dependent outcome with marker-dependent censoring. Biometrical Journal, 55(5):687-704.

Excess cumulative incidence estimation for matched survival data

C. Boschini[1,2], K. K. Andersen[1], T. H. Scheike[2]

[1] Unit of Statistics and Pharmacoepidemiology, Danish Cancer Society Research Center, Copenhagen, Denmark
[2] Department of Biostatistics, University of Copenhagen, Copenhagen, Denmark

E-mail for correspondence: crbo@sund.ku.dk

Abstract:
The excess risk of exposed groups is an important instrument in epidemiology. As discussed in our previous manuscript (Boschini et al., 2017), the excess risk analysis provides a measure of the potential impact of exposure on the public health setting. Here we extend our previous methodology to the competing risk setting, where the excess risk is described in terms of excess cumulative incidence. The model works within matched cohort studies, where individuals with a defined disease are matched with unexposed persons according to characteristics, such as country, sex and age. By exploiting the fact that the data are matched, the method is able to naturally handle two time scales, namely age and duration time. We present a regression model for the excess cumulative incidence on the lines of the binomial regression method proposed by Scheike et al. (2008). We adapt such approach with regard to the expected correlation and to the possible left truncation, given by the matched structure of the registry data. We provide an accurate description of the model by proving large sample properties of the estimators and by reporting simulation-study results. Lastly, we highlight the features of the method with an application on childhood cancer survivors.

Key words:
Cumulative incidence; Excess risk model; Generalize estimating equations; Matched cohort data; Multiple time scales.

C. Boschini, K. K. Andersen and T. H. Scheike (2017). Excess risk estimation for matched survival data. Manuscript submitted for publication.

T. H. Scheike, M. J. Zhang and T. A. Gerds (2008). Predicting cumulative incidence probability by direct binomial regression. Biometrika, 95(1):205-220.

Using multi-state modelling to facilitate informed personalised treatment planning in Follicular Lymphoma

S.E. Lacy[1], H. Wang[1], A.G. Smith[1], E. Roman[1], S. Crouch[1]

[1]University of York, York, UK

E-mail for correspondence: stuart.lacy@york.ac.uk

Abstract:
Follicular lymphoma is an incurable haematological cancer that tends to follow a remitting relapsing course; with treatment options ranging from "watch and wait" (active monitoring), chemotherapy, and radiotherapy. The treatment decision is typically dependent upon patient characteristics and disease stage. Using routine data collected by the Haematological Malignancy Research Network (www.hmrn.org), we are undertaking research into improving understanding of the treatment pathways and the impact of these decisions not only on the patient, but also on the cost to the healthcare provider. The aim is to facilitate informed decision making by providing a personalised prognostic tool that identifies a suitable treatment option at each stage of the treatment pathway, as part of a shared decision-making process between patient and clinician.
Multi-state modelling provides an appropriate toolkit for this work, owing to its ability to predict the movement of a person through a discrete state-space captured by individual-level characteristics. This work details implementation considerations of the modelling stage, including the design of the state-diagram, the fitting of transition-specific hazard functions, and the use of Discrete Event Simulation to provide estimated transition probabilities. Other considerations include how to present the results of a complex model into a deployable clinical tool.

Key words: Multi-state modelling; Follicular lymphoma; personalised treatment.

Session 5: Pseudo-observations

Chair: Priya Vart

Regression models in survival analysis with left-truncation using pseudo-observations

Mia Klinten Grand[1,2], Arthur Allignol[3], Hein Putter[2], Per Kragh Andersen[1]

[1]Copenhagen University, Copenhagen, Denmark
[2]Leiden University Medical Center, Leiden, the Netherlands
[3]Ulm University, Ulm, Germany

E-mail for correspondence: migra@sund.ku.dk

Abstract:
Pseudo-observations were introduced as a way to perform regression on a parameter related to a right-censored time-to-event outcome, such as the survival probability or the restricted mean survival time (Andersen et al., 2003). They offer an alternative to the classical method, where the relation between the covariates and the hazard is modelled. In more complex settings, such as competing risks or general multi-state models, the pseudo-observations enable the estimation of the direct effect of the covariates on e.g. the cumulative incidences or the transition probabilities, unlike the classical method.

The pseudo-observations are jackknife estimates, which represents the subjects' contribution to the non-parametric estimator of the parameter of interest. Pseudo-observations are calculated for all subjects in the sample and the regression model parameters are then obtained by solving the corresponding generalised estimating equations using the pseudo-observations as outcome.

Since the introduction of the approach there have been several extensions from the original setting. However, the proper definition and performance of pseudo-observations under left-truncation has not yet been addressed. We explored this in a simulation study, where we also looked at the performance of different versions of the pseudo-observations.

Key words: Survival analysis; Pseudo-observations; Left-truncation.

PK. Andersen and JP. Klein and S. Rosthøj (2003). Generalised linear models for correlated pseudo-observations, with applications to multi-state models. Biometrika, 90:15-27.

Goodness of fit test for regression models based on pseudo observations

Klemen Pavlič[1], Torben Martinussen[2], Per Kragh Andersen[2]

[1]Institute for Biostatistcs and Medical Informatics, Faculty of Medicine, University of Ljubljana, Ljubljana, Slovenia
[2]Section of Biostatistics, Department of Public Health, University of Copenhagen, Copenhagen, Denmark

E-mail for correspondence: klemen.pavlic@mf.uni-lj.si

Abstract:
Regression models for survival time are often formulated on the hazard scale even though hazards are not of primary interest. Relation between covariates and hazard changes when such models are transformed onto another scale (survival probability, restricted mean life time, etc). To overcome these difficulties, models based on pseudo observations were proposed. Such models allow us to directly estimate covariate effect on the mean value parameter of interest. However, they have some underlying assumptions that should be checked before valid inference could be made and results could be interpreted. We present a technique for checking the underlying assumptions for a general set of mean value parameter models. This approach is based on the cumulative sum of pseudo-residuals. The distribution of this process can be approximated by some Gaussian process and this allows us to plot it together with some realizations of this Gaussian process and visually assess the misspecification. To obtain a more objective measure, this procedure can be supplemented with a supremum test. We explore the properties of such a test for models for different mean value parameters (survival probability, restricted mean lifetime, cumulative incidence function and years lost due to a specific cause).

Key words: Goodness-of-fit; pseudo-observations; pseudo-residuals.

Modelization of the effect of covariates on dementia health indicators: approach by pseudo values

C. Sabathé[1], P. Joly[1]

[1]Univ. Bordeaux, INSERM, Bordeaux Population Health, U1219, Bordeaux - France

E-mail for correspondence: camille.sabathe@u-bordeaux.fr

Abstract:
Our objective is to study the effect of covariates on three health indicators: the probability of becoming demented, the probability of staying alive and non-demented and the restricted survival time in good health. Pseudo-values approach computed from non-parametric estimators and developed by Andersen et al. (2003) allows direct modelling of the covariate effects on these probabilities in competing risks and right censored data framework. In eldery cohort, dementia is only observed on discrete times (the diagnosis is made at visits) but the disease appears on a continuous time scale. This interval censoring makes uncertain both the exact date of the onset and the individual trajectories. Indeed, subjects who die without diagnosis of dementia will not be observed as demented. n, pseudo values are used as a variable of interest in a GLM model to see how the covariates impacts on the health indicators.
This work focus precisely on expanding the pseudo value approach to interval-censored data in order to apply them to dementia cohort. Briefly, the health indicators are estimated through penalized likelihood estimators of an illness-death model (Touraine et al. 2003).
The method is applied to the French cohort PAQUID which included more than 3,000 non-demented subjects, aged 65 years or older and followed for dementia over more than 25 years.

Key words: pseudo-values; interval-censored data; illness-death model; dementia

PK. Andersen, JP Klein and S Rosthoj (2003). Generalised Linear Models for Correlated Pseudo-Observations, with Applications to Multi-State Models. Biometrika 90 (1):15270.

C. Touraine, C. Helmer and P. Joly (2016). Predictions in an Illness-Death Model. Statistical Methods in Medical Research 25 (4):14521470.

Session 6: Dynamic prediction & competing risks

Chair: Jan Feifel

Development and validation of dynamic clinical prediction models for discrete time-to-event data with competing risks

Rachel Heyard[1], Jean-François Timsit[2] and Leonhard Held[1]

[1] Department of Biostatistics at the Epidemiology, Biostatistics and Prevention Institute, University of Zurich, Hirschengraben 84, 8001 Zurich, Switzerland
[2] UMR 1137, IAME, University Paris-Diderot, Inserm, F-75018 Paris, France

E-mail for correspondence: rachel.heyard@uzh.ch

Abstract:
Developing good prediction models is of central importance in clinical research. Techniques to perform objective Bayesian variable selection in the linear model are well developed and have been extended to the generalized linear model setting [Held et al. (2015)] and the Cox proportional hazards model [Held et al.(2016)]. We propose an extension of this methodology to discrete time-to-event data with competing risks [Heyard et al. (2017)]. In our application to a large French database, the goal is to predict the risk of a ventilator-associated pneumonia (VAP) attributed to *P. aeruginosa* in intensive care units. The competing events are extubation, death and VAP due to other bacteria. We further use a landmark approach [van Houwelingen and Putter (2012)] for dynamic Bayesian variable selection where the set of relevant predictors depends on the time already spent at risk. Additionally we determine the impact of a variable on each competing event through cause-specific variable selection. In a final step, we explain how to validate these rather complex dynamic prediction models by assessing their overall performance, calibration and discrimination.

Key words: Bayesian variable selection; competing events; discrete time-to-event model; dynamic prediction models by landmarking; cause-specific variable selection.

Leonhard Held, Daniel Sabanés Bové and Isaac Gravestock (2015). Approximate Bayesian model selection with the deviance statistic. Statistical Science, 30(2):242-257.

Leonhard Held, Isaac Gravestock and Daniel Sabanés Bové (2016). Objective Bayesian model selection for Cox regression. Statistics in Medicine, 35(29):5376-5390.

Rachel Heyard, Jean-François Timsit, Wafa Ibn Essaied and Leonhard Held (2017). Dynamic clinical prediction models for discrete time-to-event data with competing risks; a case study on the OUTCOMEREA database. Submitted.

Hans C. van Houwelingen and Hein Putter (2012). Dynamic Prediction in Clinical Survival Analysis. CRC Press.

A comparison of the beta-geometric model with landmarking for dynamic prediction of time-to-pregnancy

R. van Eekelen[1,2], H. Putter[3], M.J.C. Eijkemans[2], N. van Geloven[3]

[1]Academic Medical Center Amsterdam, the Netherlands
[2]Julius Center, Utrecht University Medical Center, the Netherlands
[3]Leiden University Medical Center, the Netherlands

E-mail for correspondence: r.vaneekelen@amc.uva.nl

Abstract:
We conducted a simulation study to compare two methods for dynamic prediction of pregnancy, i.e. predicting the probability of pregnancy over time for the same individual. The first is landmarking, a semi-parametric method introduced by van Houwelingen and Putter, where the non-parametrically estimated baseline hazard captures an increase or decrease in the probability of pregnancy over time. The second is the beta-geometric model, a parametric model introduced by Weinberg and Gladen, which is specific to applications with a discrete time-to-event. The beta-geometric model introduces a frailty, i.e. unobserved heterogeneity, according to a beta distribution and combines this with the geometric distribution. The frailty accommodates a decrease in probability of pregnancy over time due to a selection process. It is unknown what the advantages or disadvantages are of these two models and how their accuracy compares. We simulated time-to-pregnancy data according to three different true frailty distributions: beta-geometric, normal-geometric and compressed-beta-geometric, with or without a sterile subpopulation. We then compared the two methods by the following metrics: bias in average prediction, C-index and Brier score. In addition, we focus on robustness across scenarios and differences between methods in variance of these metrics between simulation replications.

Key words: Dynamic prediction; Beta-geometric; Landmarking; Cox model; Time-to-pregnancy

R. van Eekelen et al. (2017). Natural conception: repeated predictions over time. Human Reproduction, 32(2):346-53.

H. van Houwelingen and H. Putter (2012). Dynamic prediction in Clinical Survival Analysis. Chapman and Hall, 2012.

C.R. Weinberg and B.C. Gladen (1986). The beta-geometric distribution applied to comparative fecundability studies. Biometrics, 42(3):547-60.

Prediction of contralateral breast cancer using individual patient data

D.Giardiello[1,2], M.Hauptmann[1], M.J. Hooning[3], H.Oldenburg[1], M.Adank[1], A Jager[2], M.K. Schmidt[1], E.W. Steyerberg[2,3]

[1]The Netherlands Cancer Institute, Amsterdam, Netherlands
[2]Leiden University Medical Center, Leiden, Netherlands
[3]Erasmus MC,Rotterdam, Netherlands

E-mail for correspondence: d.giardiello@nki.nl

Abstract:
There are currently no validated prognostic models available to estimate risk of contralateral breast cancer (CBC) for clinical decision making. We aimed to develop and validate a prognostic model to inform patients about their CBC risk based on patient-, tumor- , and treatment characteristics of the first breast cancer. We applied the principles of individual patient data meta-analysis (IPD-MA) using 132,756 patients to estimate the risk of contralateral breast cancer (CBC) in 20 Dutch national and international studies with 4,682 CBC events and a median follow-up of 8 years . First, we imputed both sporadic and systematic missing data using a stratified multiple imputation model (SMI). Later, we developed a global Fine and Gray model stratified by study to take account into heterogeneity. Discrimination and calibration accuracy of the clinical prediction model were evaluated through an internal-external validation procedure. Results showed the calibration was reasonable, but the discrimination was limited varying between 0.51 to 0.70 across studies. However, further investigations and extensions in the current methodological statistical literature are needed both in the imputation and in validation procedures in the IPD-MA to take account into heterogeneity across studies and competing risks.

Key words: individual patient data; missing data; risk prediction; competing risks; contralateral breast cancer

I. Ahmed et al. (2014). Developing and validating risk prediction models in an individual participant data meta-analysis. BMC medical research methodology, 14(1):3.

S. Jolani et al. (2015). Imputation of systematically missing predictors in an individual participant data meta-analysis: a generalized approach using MICE. Statistics in medicine, 34(11):1841-63.

R.B. Geskus (2011). Cause-specific cumulative incidence estimation and the Fine and Gray model under both left truncation and right censoring. Biometrics, 67(1):39-49.

E.W. Steyerberg (2009). Clinical prediction models: a practical approach to development, validation, and updating. Springer Science and Business Media.

Estimating cumulative incidence functions in situations with dependent left-truncation

R. Stegherr[1], J. Beyersmann[1], C. Schaefer[2], R. Meister[2,3], A. Allignol[4]

[1]Institute of Statistics, Ulm University, Germany
[2]Charité Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universitaät zu Berlin, and Berlin Institute of Health, Pharmakovigilanzzentrum Embryotoxikologie, Institut für Klinische Pharmakoloie und Toxikologie, Berlin, Germany
[3]Beuth Hochschule für Technik - University of Applied Sciences, Berlin, Germany
[4]Merck KGaA, Darmstadt, Germany

E-mail for correspondence: regina.stegherr@uni-ulm.de

Abstract:
In a typical competing risks setting with delayed study entry the Aalen-Johansen estimator is used to estimate the cumulative incidence functions of the competing events. But one assumption for using this estimator is independent left-truncation. However, for illustrative purpose only, an example: in studies about the usage of coumarin derivatives during pregnancy the risk of experiencing an induced abortion may depend on the study entry time. This is called dependent left-truncation, since there is a dependence between left-truncation times and the times-to-event of one or possibly all of the competing events. Therefore, the Aalen-Johansen estimator can not be used and we propose a new semiparametric estimator of the marginal cumulative incidence function that does not require independence. To this end, the dependence between entry times and time-to-event is modelled using a Cox proportional hazards model and the marginal estimates are derived via inverse probability weighting arguments. Simulations as well as the data example show that the new estimator is preferable compared to the Aalen-Johansen estimator.

Key words: Dependent left-truncation; Cumulative incidence function; Cox model; Inverse probability weighting;

Joint modelling of progression-free survival and overall survival in oncology trials using the gamma threshold model

Enya Weber, Dr. Andrew Titman

E-mail for correspondence: e.weber@lancaster.ac.uk

Abstract:
In oncology trials, different clinical endpoints can be considered for the analysis of overall survival. In addition to the traditional time to death, progression-related endpoints such as progression-free survival and time to progression are often used for the evaluation of treatment effects on overall survival. A new approach[1] of modelling semi-competing risks data, where we have a terminal event such as death and non-terminal event such as progression, considers the events to arise through the first passage times of a latent gamma process. An evaluation of this modelling approach in terms of the implied transition intensities of the observable process will be presented. In addition, the modelling approach will be extended to include covariates such as treatment. An investigation of the extent to which the extended approach can increase efficiency in assessment of treatment effects on overall survival, and under what circumstances, is conducted by simulating data in a range of scenarios. This approach will be illustrated using data from a clinical trial of treatments for colon cancer.

Key words: Cancer trials; Progression-free survival; Semi-competing risks; First passage time: Gamma process.

[1] **Sildnes B, Lindqvist BH** (2017). Modeling of semi-competing risks by means of first passage times of a stochastic process. Lifetime data analysis, 1-23.

Poster session

1. Comparison of methods to integrate mediation and survival analysis

L. Burgos Ochoa [1], J. Rijnhart[1], M. Heymans [1], J. Twisk [1]

[1]VU medisch centrum (VUmc)

E-mail for correspondence: l.burgosochoa@vumc.nl

Abstract:

Statistical mediation analysis as a research goal has recently raised interest in the field of epidemiology. While mediation with continuous outcomes has been widely discussed in the literature, methods for performing mediation analysis with another type of outcomes such as survival data did not receive attention until very recently. The most common methods for modeling survival outcomes are Cox proportional hazards and accelerated failure time models. Nonetheless, it has been shown that the use of both methods incurs into a problem while performing mediation analysis with the classical mediation approach, resulting in mathematically inconsistent results. This happens due to the noncollapsibility of hazard ratios and presence of censoring in the data. To solve this issue, some methods based on the potential outcomes causality framework have been developed to estimate mediated effects with survival outcomes. This study aims to perform a Monte Carlo simulation study to evaluate and compare the performance of three different methods (Tchetgen Tchetgen, 2013; VanderWeele, 2011; Wang and Albert, 2017) developed to integrate mediation and survival analysis under conditions that reflect scenarios encountered in practice.

Key words: survival analysis; statistical mediation; Monte-Carlo simulation; potential outcomes; causal inference.

Tchetgen Tchetgen, E. J. (2013). Inverse odds ratio-weighted estimation for causal mediation analysis. Statistics in Medicine, 32(26): 4567-4580.

VanderWeele, T. J. (2011). Causal Mediation Analysis With Survival Data. Epidemiology, 22(4): 582-585.

Wang, W., and Albert, J. M. (2017). Causal mediation analysis for the Cox proportional hazards model with a smooth baseline hazard estimator. Journal of the Royal Statistical Society: Series C (Applied Statistics), 66(4): 741-757.

2. A varying-coefficient multi-state model to evaluate the effect of a longitudinal biomarker on a disease process.

Yuntao Chen[1], Douwe Postmus[1], Christine Eulenburg[1]

[1]Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

E-mail for correspondence: y.chen@umcg.nl

Abstract:

Multi-state modeling provides a framework that allows for the comprehensive analysis of a disease process. Joint modelling of longitudinal and time-to-event data has been proposed to investigate the association between a longitudinal biomarker and the transition intensities in a multi-state model. However, it can be very difficult to specify an appropriate model for the longitudinal biomarker if it varies greatly among patients. It may be more reasonable to consider that the effect of the biomarker on a transition intensity has less variation among different patients at a specific time point than the biomarker itself. In this context, under the framework of multi-state models, we incorporate the longitudinal biomarker with the counting process method and construct a smooth function for the effect of the biomarker on the transition intensities. In a simulation study, a calibration method is used to compare the predictive performance of our model with that of a joint model in different scenarios: 1) violation of the proportional hazard assumption; 2) misspecification of the biomarker trajectories; 3) misspecification of the dependence function. Finally the advantages and disadvantages of the two modeling approaches will be illustrated in an example in heart failure disease progression.

Key words: Multi-state model, Counting process, Smooth function, Joint modelling, Calibration.

3. Multiple imputation strategies for interval-censored event times

E. Curnow[1,2], K. Birnie[2], M. T. May[2], K. M. Tilling[2]

[1]NHS Blood and Transplant, Bristol, UK
[2]Department of Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

E-mail for correspondence: elinor.curnow@bristol.ac.uk

Abstract:

Graft-versus-host disease (GvHD) is a major cause of mortality after haematopoietic stem cell transplant (HSCT). Where GvHD has occurred, the exact date of onset is often not reported. This means that the GvHD time is interval-censored, with the censored interval spanning the patients entire follow-up period. In this setting an expectation-maximisation approach is problematic. Multiple imputation (MI) is an alternative procedure for handling interval-censored data. After imputation, analysis can be performed using any complete data method. This is helpful in HSCT studies where competing risks and multiple events are common.

We use a simulation study to compare a range of MI strategies for handling the interval-censored times. Data are generated from a cohort of patients who received HSCTs between 1996 and 2015, from cord blood donated to the NHS Cord Blood Bank, UK. We describe the impact on the distribution and overall estimate of the cumulative incidence of GvHD and the width of the associated confidence intervals.

Multiple imputation of interval-censored event times allows more accurate estimation of cumulative incidence of GvHD. There is a need for more sophisticated missing data methods to be adopted by the HSCT community, which will require them to be implemented in standard software.

Key words: interval censoring; multiple imputation; competing risks; transplantation.

4. Flexible Bayesian Additive Joint Models for Longitudinal and Time-to-Event Data: Applicacion to Liver Transplantation Data

Carla Díaz-Louzao[1], Francisco Gude[2], Carmen Cadarso-Suárez[1]

[1]Unit of Biostatistics, Department of Statistics, Mathematical Analysis and Optimization, University of Santiago de Compostela, Spain
[2]Clinical Epidemiology Unit, Hospital Clínico, Santiago de Compostela, Spain

E-mail for correspondence: carla.diaz.louzao@usc.es

Abstract:

In patients who underwent liver transplantation, it is important to assess what factors could predict the risk for death after this procedure. Age, sex, body mass index, alcohol intake and parenteral nutrition, among others, were analyzed. Glucose levels just before the surgery and once a day during the following week were also included. In order to study the risk of death of this patients, joint modelling approaches were used. This procedure is an appropriate way to study the relationship between the longitudinal and the time-to-event processes.

In Orthopedic Liver Transplantation data, both longitudinal trend and the risk for death show non-linear profiles. This, together with the fact that the relationship between the biomarker and the survival process may vary over time, has led us to implement the methodology presented in Köhler et al. (2017), which consists of a new approach to the problem from the Bayesian point of view.

Key words: Joint modelling, longitudinal, time-to-event, transplantation.

Acknowledgements: This study was supported by Grant from the Program of Aid to the Predoctoral Stage of the Xunta de Galicia and European Social Fund 2014/2020.

Köhler M, Umlauf N, Beyerlein A, Winkler C, Ziegler A, Greven S. (2017). Flexible Bayesian additive joint models with an application to type 1 diabetes research. Biometrical Journal, 00:1-18.

5. P-splines and Laplace approximations for fast Bayesian inference in a flexible promotion time cure model

Gressani Oswaldo [1], Philippe Lambert [1,2]

[1]Institute of Statistics, Biostatistics and Actuarial Sciences, Université catholique de Louvain, Louvain-la-Neuve, Belgium.
[2]Faculté des Sciences Sociales, Université de Liège, Liège, Belgium.

E-mail for correspondence: oswaldo.gressani@uclouvain.be

Abstract:
In traditional survival models, it is assumed that the event of interest will occur for all subjects under study, provided the follow-up is sufficiently extended in time. To relax this assumption, a class of models coined "cure rate models" have emerged in the literature, allowing situations where a fraction of subjects may never develop the monitored event. We propose a novel approximate Bayesian methodology for inference in a promotion time cure model, by exploiting the synergy between P-splines and Laplace approximations. P-splines provide a flexible smoothing of the survival quantities of interest, while Laplace approximation is an elegant tool for obtaining fast and accurate approximations of posterior distributions of latent model variables. An attractive feature of this approach is that point estimators and credible intervals can be straightforwardly constructed, even when considering non-trivial functionals of latent variables. The properties of the proposed methodology are evaluated using simulations and illustrated on oropharynx carcinoma data. This flexible and sampling-free approach can be considered a serious competitor of Markov chain Monte Carlo for Bayesian inference in cure models.

Key words: Promotion time cure model; P-splines; Laplace approximation; Bayesian inference.

Eilers, P.H.C. and Marx, B.D. (1996). Flexible smoothing with B-splines and penalties. Statistical Science, 11(2):89-102.

Bremhorst, V. and Lambert, P. (2016). Flexible estimation in cure survival models using Bayesian P-splines. Computational Statistics and Data Analysis, 93:270-284.

Rue, H., Martino, S. and Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using Integrated Nested Laplace Approximations. Journal of the Royal Statistical Society (Series B), 71(2):319-392.

Lang, S. and Brezger, A. (2004). Bayesian P-splines. Journal of Computational and Graphical Statistics, 13(1):183-212.

6. Restricted mean lifetime estimation in the presence of competing risks from within the flexible parametric modelling framework

Sarwar Islam Mozumder[1], Mark J. Rutherford[1], Paul C. Lambert[1,2]

[1]Biostatistics Research Group, Department of Health Sciences, Centre of Life Sciences, University of Leicester, Leicester, UK
[2]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

E-mail for correspondence: si113@le.ac.uk

Abstract: As more detailed cancer registry data becomes available, models on cancer survival are increasing in complexity. It is therefore of greater importance to make available simple and interpretable measures that can be presented to patients or clinicians after fitting such models. In this talk, the restricted mean lifetime estimate is presented as one such measure. This quantifies the expected number of life-years lost at time t following death from a particular cause. Estimation is made in the presence of competing risks within the flexible parametric modelling framework. One has the choice of either directly modelling the cause-specific cumulative incidence function, or by calculating the cause-specific cumulative incidence function using individual models for all the cause-specific hazards. Estimation of restricted mean lifetimes after fitting models under both approaches are demonstrated. Numerical integration from t_0 to t_1 by way of the Gauss-Legendre quadrature method is used for obtaining the expected number of life-years lost before time t_1 due to cause k and confidence intervals are provided using the delta method. Finally, an extension of the user-friendly command, `stpm2cr`, in Stata is presented as an implementation of the methods using SEER public-use colorectal data.

Key words: Competing risks, Restricted mean, Life-years lost, Flexible parametric, Gaussian quadrature.

P.K. Andersen (2013). Decomposition of number of life years lost according to causes of death. Statistics in medicine, 32(30):5278-5285.

S.R. Hinchliffe and P.C. Lambert (2013). Flexible parametric modelling of cause-specific hazards to estimate cumulative incidence functions. BMC medical research methodology, 13(13):1471-2288.

S.I. Mozumder, M.J. Rutherford and P.C. Lambert (2017). Direct likelihood inference on the cause-specific cumulative incidence function: A flexible parametric regression modelling approach. Statistics in medicine, 1-16.

7. Extreme value estimation with Cox model

Kévin Jaunâtre[1], Ion Grama[1]

[1]Université de Bretagne Sud

E-mail for correspondence: kevin.jaunatre@univ-ubs.fr

Abstract:

We propose an extension of the regular Cox's proportional hazard model which allows to estimate probabilities of rare events. It is known that when the data are heavily censored at the upper end of the distribution, the estimation of the tail of the distribution is not reliable. To allow an estimation of the distribution beyond the last observed data, we suppose that the survival data are in the domain of attraction of the Fréchet distribution conditionally to covariates. Under this condition, by the Fisher-Tippett-Gnedenko theorem, the tail of the baseline distribution can be adjusted by a Pareto distribution with parameter θ beyond a threshold τ. The survival distributions conditioned to the covariates are easily computed from the baseline. We also propose an aggregated estimate of the survival probabilities. A procedure allowing an automatic choice of the threshold is presented and an application on a real data set collected from a project ongoing in France is given.

Key words: Cox model; Extreme value theory; Pareto distribution; Survival probabilities

8. Dependent competing risks in a mixture cure model

Radoslav Kovar[1]

[1]Department of Statistics and Probability, Faculty of Informatics and Statistics, University of Economics, Prague, Czech Republic

E-mail for correspondence: xkovr11@vse.cz

Abstract:

Survival models of competing risks, survival models with a cure fraction and combinations of both have been applied in various fields. These models are often based on the assumption of independence among the competing survival times. Violation of this assumption leads to biased parameter estimates and biased estimates of their standard errors. This paper therefore explores the size of these biases within a model containing a bivariate parametric copula, two parametric marginal survival functions, a cure fraction and a dummy covariate. Necessary and sufficient conditions of the correctly specified model and of the model with incorrectly assumed independent risks are derived. The differences between both models are evaluated in a simulation study with a varying degree of dependence and for different parametrisations of the two marginal survival functions. The results are intended to help during sensitivity analyses where the structure of dependence is unknown.

Key words: Competing risks; Cure model; Dependent censoring; Copula functions; EM algorithm.

9. Target, Anticipated, and Regulated Received Dose Intensity

Three ways of looking into osteosarcoma survivorship

C. Lancia[1]

[1]Leiden University

E-mail for correspondence: c.lancia@math.leidenuniv.nl

Abstract:

Background

Received dose intensity (RDI), a keystone concept in chemotherapy, is the amount of cytostatic agents delivered in a time-window (Hryniuk and Bush, 2984; Hryniuk and Goodyear, 1990). In vitro models support the hypothesis that higher RDI achieves higher tumorous-cell-kill rate and better survival (Foote, 1998). Yet, the role of dose-intensity role in cancer survivorship is still unclear for many cancer types, including osteosarcoma (Ozols, 2007; Citron, 1999; Smeland et al., 1999; Lewis et al., 2007).

Patients and Methods

Data come from randomised clinical trial MRC BO06 (ISRCTN86294690; Lewis et al., 2007). We compute RDI in three different ways, form patient strata accordingly, and contrast the different estimation of event-free survival in a landmark analysis at 180 days since registration. We compare target RDI (calculated at the beginning of treatment on the allocated regimen), achieved RDI (individually computed on actual given dose and treatment duration) and regulated RDI (which extends the achieved calculation over time).

Results

Stratifying patients by target or achieved RDI shows survival curves that are practically equivalent. On the contrary, regulated RDI yields a completely different picture and produces strata that are markedly different from the point of view of survival.

Conclusions

This work demonstrates the benefits of addressing the relationship between received dose intensity and survival at both individual and longitudinal level. The work shows also that the concept of RDI might be an inadequate risk factor for osteosarcoma.

Key words: Cancer treatment; Clinical trials.

W.M. Hryniuk and H. Bush (1984). The importance of dose intensity in chemotherapy of metastatic breast cancer. J Clin Oncol, 2(11):1281–1288.

W.M. Hryniuk and M. Goodyear (1990). The calculation of received dose intensity. J Clin Oncol, 8(12):1935–1937.

M. Foote (1998). The importance of planned dose of chemotherapy on time: do we need to change our clinical practice?. The Oncologist, 3(5):365–368.

R.F. Ozols (2007). Ovarian cancer: Is dose intensity dead? J Clin Oncol, 25(27):4157–4158.

M.L. Citron (1999). Dose density in adjuvant chemotherapy for breast cancer. Cancer Invest. 2004; 22(4):555–568.

S. Smeland, C. Muller, T.A. Alvegard, et al. (1999). SSGVIII study: prognostic factors for outcome and role of replacement salvage chemotherapy for poor histologic responders. Eur J Cancer, 39: 488–494.

I.J. Lewis, M.A. Nooij, J. Whelan, et al. (1999). Improvement in histologic response but not survival in osteosarcoma patients treated with intensified chemotherapy: a randomized phase III trial of the European Osteosarcoma Intergroup. J Natl Cancer Inst, 99(2):112–128.

10. A comparison of methods to account for time-varying treatment use when developing a prognostic model

R. Pajouheshnia[1], N. Schuster[1], R.H.H. Groenwold[1,2], L.M. Peelen[1]

[1]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands
[2]Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands

E-mail for correspondence: R.Pajouheshnia@umcutrecht.nl

Abstract:

Prognostic models used in deciding whether to initiate treatment need to estimate a patients outcome risk in the absence of treatment, but data used to develop prognostic models typically includes treated individuals. We compare methods to account for time-varying treatment use in prognostic model development. Cox regression was used to develop models to predict 3-year mortality risk without beta-blocker use in COPD patients, using electronic health record data from 2095 patients (287 events, median follow-up=3 years). Models were developed following six approaches: (i) ignore treatment, (ii) exclude beta-blocker users at baseline, (iii) exclude beta-blocker users at baseline and during follow-up, (iv) ii+ include treatment during follow-up as a binary covariate, (v) include beta-blocker use as a time-varying covariate, (vi) censor treated individuals, (vii) vi+ inverse probability of censoring weighting, (viii) inverse probability of treatment weighting, and were compared in terms of discrimination (c-statistic), calibration (slope, plots), accuracy (Brier score) and reclassification in the full cohort and the untreated subset. Compared to a model that ignored treatment use (full data: c-statistic=0.80 , Brier=0.058), we found no differences in performance across methods (ii)-(iv). The results of (v)-(viii) will be presented at the conference. In conclusion, we found no difference between methods to account for the effects of time-varying treatments. Simulation studies will be conducted to confirm these empirical findings.

Key words: Prognosis; Prediction; Treatment drop-in; Time-varying covariates; Inverse probability weighting.

11. Penalized Variable Selection in Accelerated Failure Time Models with Random Effects

Eunyoung Park[1], Il Do Ha[1]

[1]Department of Statistics, Pukyong National University, Busan 48513, South Korea

E-mail for correspondence: 951025park@gmail.com

Abstract:

Accelerated failure time (AFT) model is a linear model under the log-transformation of survival time and has been introduced as an useful alternative of proportional hazards (PH) model. In this talk we propose a variable selection procedure of fixed effects in the AFT model with random effects using a penalized hierarchical likelihood (h-likelihood) approach (Ha et al., 2014, 2018). Here we use two popular penalty functions, LASSO and SCAD. The performance of the proposed method is evaluated via simulation studies under setting of clustered survival data. The proposed method is illustrated with two practical data, including the kidney infection data and bladder cancer recurrence data from multi-center clinical trial. Furthermore, we discuss about an extension of the proposed method to competing risks.

Key words: AFT model; H-likelihood; Penalized likelihood; Random effects; Variable selection.

Ha, I. D., Jeong, J.-H. and Lee, Y. (2018). *Statistical modelling of survival data with random effects: h-likelihood approach*. Springer, in press.

Ha, I. D., Pan, J., Oh. S. and Lee, Y. (2014). Variable selection in general frailty models using penalized h-likelihood. Journal of Computational Graphical Statistics, 23:1044-1060.

12. Novel statistical methods for efficient identification of biomarkers for personalized cancer treatment

D.A. Stram[1], K. Jozwiak[1], S.C. Linn[2], M. Hauptmann[1]

[1]Division of Psychosocial Research and Epidemiology, Netherlands Cancer Institute, Amsterdam
[2]Division of Molecular Pathology, Netherlands Cancer Institute, Amsterdam

E-mail for correspondence: d.stram@nki.nl

Abstract:

Identification of predictive biomarkers, i.e. markers that can be used to identify patients who are most likely to benefit from a specific treatment, has proven difficult, partially due to problems relating to statistical power and financial cost. Case-only methods have been suggested as a potential way of reducing costs while maintaining the statistical power of preventive biomarker studies . Such methods provide an estimate of marker-treatment interaction equivalent to that from standard time-to-event analysis, under certain assumptions. We apply various case-only methods and the standard Cox regression interaction analysis to four previously conducted predictive biomarker studies, including data from two randomized trials, on breast cancer survival and the BRCA1-like marker in combination with high dose chemotherapy as well as estrogen receptor phosphorylation in combination with tamoxifen treatment. In addition, we perform simulation studies to evaluate the robustness of case-only methods with respect to, among others, violations of the rare disease assumption and independence between marker and treatment. Ultimately, we will characterize the circumstances where these case-only methods can be used effectively, without bias or loss of power relative to standard full cohort analyses, and which methods are most appropriate in such scenarios.

Key words: predictive biomarkers; cancer; personalized medicine; case-only analysis; simulation study.

13. Sparse Variable selection for competing risks

Leili Tapak[1], Omid Hamidi[2]

[1]Department of Biostatistics and Epidemiology, School of Public Health and Modeling of Noncommunicable Diseases Research Center, Hamadan University of Medical Sciences, Hamadan 65175-4171, Iran
[2] Department of Science, Hamedan University of Technology, Hamedan, 65155, Iran

E-mail for correspondence: Leilitapak@gmail.com

Abstract:

Abstract text: Sparse variable selection has become an increasingly interesting topic for many medical research especially in high-dimension setting. There are found extensive studies for standard time-to-event data that can not be applied for competing risks setting directly. One of the popular modeling approaches is proportional subdistribution hazards model that was proposed by Fine and Gray and it directly models the cumulative incidence function. Shrinkage techniques has been shown to have acceptable performance in variable selection based on different regression models including Cox PH model. The aim of the present study was to investigate the performance of two selection methods developed in the framework of variable selection and competing risks. The adaptive LASSO, SCAD, SICA, adaptive elastic net and minimax concave penalty (MCP) were used selected as the variable selection methods. The stability of the selected variables subset by each method was investigated using a HIV/AIDS data set.

Key words: Competing risks;Variable selection; HIV/AIDS.

Zhixuan Fu1, Chirag R. Parikh2 and Bingqing Zhou (2016). Penalized variable selection in competing risks regression. Lifetime Data Anal.

14. Competing Risks in Pregnancy Prognosis: A study on the effect of delivery as a competing event for antenatal adverse events in women with early-onset pre-eclampsia.

Lucy Teece[1], Dr Kym Snell[1], Prof Danielle van der Windt[1], Dr Sally Wilkes[2], Shakila Thangaratinam[3], Prof Richard Riley[1]

[1]Institute of Primary Care and Health Sciences, Keele University.
[2]School of Medicine, University of Nottingham.
[3]Multidisciplinary Evidence Synthesis Hub, Queen Mary University of London.

E-mail for correspondence: l.teece@keele.ac.uk

Abstract:

Early-onset pre-eclampsia contributes to a number of serious pregnancy complications. The associated clinical deterioration necessitates intensive care in one third of women. The only known cure for the condition and the impending complications is inducing delivery, which carries risks for the pre-term baby.

Accurate estimation of the risks of adverse events in women diagnosed with pre-eclampsia is required to plan appropriate management and prioritisation of high-risk women. Existing models for early-onset pre-eclampsia (Thangaratinam 2017) have incorporated the effect of treatment via early delivery through the prediction of a combined outcome, adverse maternal outcome or pre-term delivery, which may have limited clinical use. As delivery of the baby prevents the occurrence of an antenatal adverse event, delivery may be seen as a competing risk and could be modelled as such.

Flexible parametric prognostic models (Royston 2002) were developed using data collected for the Prediction of Risks in Early Onset of Pre-Eclampsia study. Two models were developed, one using standard survival methods to predict the net risk of antenatal adverse events, the second using competing risks methods (Lambert 2017) to predict the risk of antenatal adverse events accounting for pre-term delivery as a competing event. The prognostic performance of the two models was examined using an external dataset.

Key words: Competing; Risks; Prognosis; Flexible; Parametric.

Thangaratinam, Shakila, et al. (2017). Prediction of complications in early-onset pre-eclampsia (PREP): development and external multinational validation of prognostic models. BMC medicine, 15.1:68.

Royston, Patrick, and Mahesh KB Parmar (2002). Flexible parametric proportionalhazards and proportionalodds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. Statistics in medicine, 21.15:2175-2197.

Lambert, Paul C., et al. (2017). Flexible parametric modelling of the causespecific cumulative incidence function. Statistics in medicine, 36.9:1429-1446.

15. Multi-state models to evaluate the benefit of a surgical procedure

Guillermo Villacampa[1], Guadalupe Gómez[2], Montse Rué[3], Rodrigo Dienstmann[1], Neda Stjepanovic[1]

[1]Vall d'Hebron Institute of Oncology
[2]Universitat Politècnica de Catalunya
[2]Universitat de Lleida-IRBLleida

E-mail for correspondence: gvillacampa@vhio.net

Abstract:

Data with multiple endpoints and/or intermediate events is increasingly common in biostatistics. In this context, statistical models that go beyond classical survival analysis are needed. Multi-state models (MM) allow a more flexible approach to model complex data in clinical research.

The present study uses MM to estimate the benefit of bilateral risk-reducing salpingo-oophorectomy (RRSO) as a surgical procedure to reduce the risk of breast cancer in women with germline BRCA 1/2 mutation. The benefit of this intervention was put into question following the detection of biases in previous studies, such as immortal time bias and informative censoring. The study population includes a retrospective cohort of 430 women from four Spanish hospitals.

After comparing with the standard Cox model and a Cox extended model with time dependent variables, a Markov MM with two transient states and two absorbing states has been considered the best alternative to account for potential biases and to provide an accurate estimation of the effect of RRSO. Additionally, this model has allowed us to estimate the impact of different covariables in each transition and the cumulative incidence of the relevant events.

In conclusion, MM appears to be the most adequate approach to study survival data with multiple endpoints.

Key words: Multi-state models 1; Survival analysis 2; Immortal time bias 3; Extended Cox Model 4; Breast cancer 5.

16. Estimation of the population-attributable fraction for survival settings with time-dependent exposures

M. K. von Cube[1], M. Schumacher[1], M. Wolkewitz[1]

[1]Institute for Medical Biometry and Statistics, Faculty of Medicine and Medical Center, University of Freiburg, Freiburg, Germany

E-mail for correspondence: cube@imbi.uni-freiburg.de

Abstract:

There is an increasing interest in understanding the impact and consequences of nosocomial infections (NIs) for patients in intensive-care. The population-benefit from the extinction of NIs is expressed by the population-attributable fraction (PAF). It relates the overall risk of an outcome (e.g. death) to the risk of the outcome among unexposed patients.

Difficulties in the estimation of the PAF arise due to the fact that the occurrence of NIs is a time-dynamic process. Moreover, adjustment for time-dependent confounders is essential to obtain an unbiased estimator. Little literature is available accommodating an estimation of the PAF in a data setting where both exposure and outcome are time-dependent.

We propose dynamic prediction by landmarking to estimate a PAF in this data situation. At each landmark the problem is reduced to a time-independent setting. Then, estimation is simply performed by using a generalized-linear model accounting for the current exposure state and further covariates. The method is explored in a simulation study and applied to a large French database of intensive-care unit patients to estimate the population-benefit of a pathogen-specific intervention that could cure ventilator-associated pneumonia caused by the pathogen *Pseudomonas aeruginosa*.

Key words: attributable risk; internal time-dependent exposure; competing risks; multi-state models; hospital-acquired infections.

H. van Houwelingen and H. Putter (2011). Dynamic prediction in clinical survival analysis. CRC Press.

M. Bekaert and S. Vansteelandt and K. Mertens (2010). Adjusting for time-varying confounding in the subdistribution analysis of a competing risk. Lifetime data analysis, 1(16):45–70.

M. Schumacher M. and Wangler and M. Wolkewitz and J. Beyersmann (2007). Attributable mortality due to nosocomial infections. A simple and useful application of multistate models. Methods of information in medicine, 5(46):595.

17. Accounting for Length of Stay in Regression Models in Hospital Epidemiology

S. Weber[1], M. Wolkewitz[1]

[1]Institute for Medical Biometry and Statistics, Faculty of Medicine and Medical Center – University of Freiburg, Freiburg, Germany

E-mail for correspondence: sweber@imbi.uni-freiburg.de

Abstract:

Performing risk-factor analysis of hospital acquired infections (HAI) researchers are focusing different challenges, such as time dependency and competing risks, see Wolkewitz et al. (2014). Time at risk has an impact on acquiring HAI but there are different approaches how to adjust for time at risk in regression models. In the literature there are several examples investigating the risk of HAI by considering odds ratios adjusted for length of stay or time at risk in order to model HAI via time. (E.g. we have found recent examples from highly ranked journals in hospital epidemiology.)

We consider the competing risk situation with death and discharge without HAI as competing risks of HAI as well as a multi-state model with HAI as an intermediate event.

We discuss the approach of time adjusted logistic regression and whether it is suitable in this setting. We compare this approach with appropriate analyses incorporating competing risks via the cumulative incidence function, hazard ratios and subdistribution hazard ratios, being considered as reference measures. Therefore a real data example is analysed and a simulation study is performed. Furthermore we discuss time adjusted logistic regression from a mathematical point of view.

Key words: Hospital acquired Infection; Competing Risks; Length of Stay; Time at Risk; Risk-Factor Analysis

M. Wolkewitz et al. (2014). Interpreting and comparing risks in the presence of competing events. BMJ, 349:g5060-g5060.

Programme Details

The schedules for the three conference days are given below. In between all sessions there will be breaks with refreshments. Abstracts of all presentations can be found in the abstract section on page 17.

Tuesday April 24th	
13.00 - 13.30	Registration
13.30 - 13.45	Welcome by Organising Commitee
13.45 - 14.30	Tutorial on Multistate Models by Prof.dr. Hein Putter - Part 1
14.45 - 15.45	Tutorial on Multistate Models by Prof.dr. Hein Putter - Part 2
16.00 - 16.45	Tutorial on Multistate Models by Prof.dr. Hein Putter - Part 3
17.00 - 18.30	Welcome reception at Leiden City Hall (Stadhuisplein 1)
19.00 - 22.00	Optional Dinner at Restaurant La Bota (Herensteeg 9)

Wednesday April 25th	
09.15 - 09.30	Registration
09.30 - 09.45	Welcome by Organising Committee
09.45 - 10.30	**Session 1: Causality**
	Chair: Cristina Boschini
	Johan Steen:
	A simple to implement approach for estimating attributable in-hospital mortality of a time-varying exposure
	Camila Olarte Parra:
	Estimating the impact on survival of immediate vs delayed kidney transplantation
	Anne Helby Petersen:
	Sibling models in the Cox framework: A panacea to inferring causality from observational data or a new source of problems?
10.45 - 11.30	**Session 2: Informative censoring**
	Chair: Regina Stegherr

	Nicole Barthel:
	Modeling recurrent event times subject to right-censoring with D-vine copulas
	Alessandro Gasparini:
	Joint models for survival and longitudinal data when the observation process is informative
	Gerrit Toenges:
	Drop-outs in studies with a recurrent event endpoint: Consequences of violating the independent censoring assumption
11.45 - 12.30	**Session 3: Flexible parametric modelling**
	Chair: Stuart Lacy
	Sarah Booth:
	Building and Validating Flexible Parametric Survival Models
	Priya Vart:
	Estimating age-at-onset distribution of the asymptomatic stage of a disease
	Dimitra Kleio Kipourou:
	Estimating cumulative incidences using flexible parametric models at individual and population levels
12.30 - 13.30	Lunch
13.30 - 14.30	Keynote speaker: Dr. Maja Pohar Perme
14.55 - 16.10	**Session 4: Multiple time scales** & **multistate models**
	Chair: Camille Sabathé
	Jan Feifel:
	A Non-standard Case-Control Design With A View Towards Rare Time-dependent Exposures in Hospital Epidemiology
	Davide Paolo Bernasconi:
	Modelling the impact of time to the intermediate event in the illness-death model
	Kassu Mehari Beyene:
	Time-dependent ROC Curve Estimation with Cure Fraction
	Cristina Boschini:
	Excess cumulative incidence estimation for matched survival data
	Stuart Lacy:
	Using multi-state modelling to facilitate informed personalised treatment planning in Follicular Lymphoma
16.10 - 17.20	**Poster Session**

19.00 - 23.00	Conference dinner at Restaurant Waag (Aalmarkt 21)

Thursday April 26th	
09.30 - 10.30	Keynote speaker: Dr. Birgit Lissenberg-Witte
10.30 - 11.15	**Session 5: Pseudo-observations**
	Chair: Theodor Balan
	Mia Klinten Grand:
	Regression models in survival analysis with left-truncation using pseudo-observations
	Klemen Pavlič:
	Goodness of fit test for regression models based on pseudo observations
	Camille Sabathé:
	Modelization of the effect of covariates on dementia health indicators: approach by pseudo values
11.30 - 12.45	**Session 6: Dynamic prediction & competing risks**
	Chair: Jan Feifel
	Rachel Heyard:
	Development and validation of dynamic clinical prediction models for discrete time-to-event data with competing risks
	Rik van Eekelen:
	A comparison of the beta-geometric model with landmarking for dynamic prediction of time-to-pregnancy
	Daniele Giardiello:
	Prediction of contralateral breast cancer using individual patient data
	Regina Stegherr:
	Estimating cumulative incidence functions in situations with dependent left-truncation
	Enya Weber:
	Joint modelling of progression-free survival and overall survival in oncology trials using the gamma threshold model
12.45 - 13.00	Closing remarks, awards announcements
13.00 - 14.00	Lunch