UNIVERSITÄT BONN

ukb universitäts klinikum bonn

imbie

10th Survival Analysis

for Junior Researchers Conference

19th - 21st March 2025

University Club Bonn

Conference Book

abbvie

I B S

Roche

ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS

Springer

Contents

Welcome

Welcome to the *Survival Analysis for Junior Researchers (SAfJR)* 2025 Conference in Bonn!

We are excited to present this conference book for the SAfJR 2025 conference. Over the next three days in Bonn, we look forward to insightful discussions, innovative research, and valuable networking opportunities.

This year marks the 10th edition of the SAfJR conference, continuing its tradition of success in different European cities. As before, this year's conference unites early-career statisticians specializing in time-to-event data, offering a unique platform for scientific discussions, practical applications, and the opportunity to present research while connecting with fellow junior researchers.

We sincerely appreciate the many valuable contributions that have made the 2025 conference possible. Thanks to these submissions, we were able to curate an engaging program covering a wide range of topics, including Epidemiology, Dynamic prediction models, Pseudo-observations, High-dimensional survival analysis and machine learning, Cure models, Causality, Pharmaceutical statistics and clinical trials, Parametric regression models, and Competing risks and multistate models.

This year, SAfJR is honored to feature three esteemed experts in the field of survival analysis. Morten Overgaard (Aarhus, Denmark) will present on regression analysis using jack-knife pseudo-observations, Nan van Geloven (Leiden, The Netherlands) will share expertise on causal prediction of time-to-event outcomes, and Dennis Dobler (Dortmund, Germany) will discuss resampling methods in survival and event history analysis. We look forward to their contributions, which will provide valuable insights and perspectives on these highly relevant topics.

The scientific program will be complemented by numerous opportunities for interaction. Regular breaks, shared lunches, a poster session, and a social program, including the conference dinner and the artificial intelligence exhibition at the Deutsches Museum ("German museum"), will create plenty of opportunities to network and exchange ideas. We hope the conference will inspire you, will foster new connections, and will leave you with a memorable experience.

If you have any questions, please feel free to reach out to us directly at the conference venue or contact us at safjr2025@imbie.uni-bonn.de. We are happy to assist you!

Looking forward to a successful and engaging event - your SAfJR 2025 organizing team
from the Institute for Medical Biometry, Informatics and Epidemiology (IMBIE)

Alina Schenk, David Köhler, Jochen Kamuf, Matthias Schmid, Tanja Hoffmann, and Vanessa Basten

Program

Wednesday 19th March 2025

08:00 - 09:00		Registration
09:00 - 09:15		Welcome
09:15 - 10:15		**Session 1 - Epidemiology** Chair: Benjamin Aretz
	Yangfan Li Oxford, United Kingdom	Risk prediction using case-cohort samples: A scoping review and empirical comparison
	Judith Vilsmeier Ulm, Germany	Implication of the choice of time scales in survival analysis
	Bor Vratanar Ljubljana, Slovenia	Leveraging cancer incidence for lead time estimation in cancer screening programmes
10:15 - 10:35		Coffee Break
10:35 - 11:35		**Session 2 - Dynamic prediction models** Chair: Anders Munch
	Niklas Hagemann Cologne, Germany	Capturing subgroup-specific time-variation in covariate effects in Cox-type hazard regression models
	Pedro Miranda Afonso Rotterdam, The Netherlands	Dynamic prediction of survival benefit to inform liver transplant decisions in hepatocellular carcinoma
	Mirko Signorelli Leiden, The Netherlands	Dynamic prediction with numerous longitudinal predictors: How to combine the best of both worlds (landmarking and joint modelling) through penalized regression calibration
11:35 - 11:45		Short Break
11:45 - 12:45		**Keynote** Chair: Alina Schenk
	Morten Overgaard Aarhus, Denmark	Regression analysis with jack-knife pseudo-observations
12:45 - 13:45		Lunch Break
13:45 - 14:25		**Session 3 - Pseudo-observations** Chair: Alina Schenk
	Simon Mack Dortmund, Germany	Bootstrap-based inference for pseudo-value regression models
	Nickson Murunga Leicester, United Kingdom	Implications of pseudo-observations in prognostic modelling: Addressing left truncation
14:25 - 18:00		Excursion: Mission AI, Deutsches Museum
18:00 - 20:00		**Poster Session**

Thursday 20th March 2025

08:30 - 9:50		**Session 4 - High-dimensional survival analysis and machine learning** Chair: Thomas Welchowski
	Antoine Caillebotte Paris, France	Estimation and variables selection in a joint model of survival times and longitudinal data with random effects
	Riccardo De Santis Siena, Italy	Sign-flip test for coefficients in the Cox regression model
	Anders Munch Copenhagen, Denmark	Targeted learning with right-censored data using the state learner
	Simon Wiegrebe Munich, Germany	Deep learning for survival analysis: A review
09:50 - 10:15		Coffee Break
10:15 - 11:35		**Session 5 - Cure models** Chair: Marta Cipriani
	Morine Delhelle Ottignies-Louvain-la-Neuve, Belgium	Copula based dependent censoring in cure models with covariates
	Blanca E. Monroy-Castillo A Coruña, Spain	Testing the effect of multiple covariates on cure rates in mixture cure models based on distance correlation
	Beatriz Piñeiro-Lamas A Coruña, Spain	The sicure R package: Single-index mixture cure models
	Tsz Pang Yuen Amsterdam, The Netherlands	Testing for sufficient follow-up in survival data with covariates
11:35 - 11:45		Short Break
11:45 - 12:45		**Keynote** Chair: Matthias Schmid
	Nan van Geloven Leiden, The Netherlands	Causal prediction of time-to-event outcomes
12:45 - 13:45		Lunch Break

13:45 - 15:05		**Session 6 - Causality** Chair: Yujun Xu
	Niklas Maltzahn Oslo, Norway	Robust estimation of occupation probabilities of latent multi-state processes
	Ilaria Prosepe Leiden, The Netherlands	Interventional dynamic updating of prognostic survival models in a pandemic environment
	Alice Marion Richardson Canberra, Australia	Surviving your PhD: An analysis of time to completion data
	Sandra Schmeller Ulm, Germany	A ”what if”-interpretation of the Kaplan-Meier estimator and, in general, no such interpretation for competing risks
15:05 - 15:35		Coffee Break
15:35 - 16:55		**Session 7 - Pharmaceutical statistics and clinical trials** Chair: Sam Doerken
	Lucia Ameis Cologne, Germany	A non-parametric proportional risk model to assess a treatment effect in an application to randomized controlled trials
	Moritz Fabian Danzer Münster, Germany	Exhausting the type I error level in a group-sequential design with a closed testing procedure for progression-free and overall survival
	Beatriz Farah Paris, France	Sample size calculation based on differences of quantiles from right-censored data
	Chloé Szurewsky Paris, France	One-sample survival tests for non-proportional hazards in oncology clinical trials: A simulation study
16:55 - 19:00		Evening Break
19:00 - 22:00		**Conference dinner at Godesburg Castle**

Friday 21st March 2025

08:30 - 09:50		**Session 8 - Parametric regression models** Chair: David Köhler
	Antoniya Dineva Bielefeld, Germany	A “double copula” model for semi-competing risks data
	Gilbert Kiprotich Munich, Germany	Incorporation of a mixture distribution on frailty regression model for clustered survival data
	Marilena Müller Heidelberg, Germany	Comparing a time-to-event endpoint in a two-arm trial investigating personalized treatment
	Thomas Welchowski Zurich, Switzerland	R-package discSurv: A toolbox for discrete time survival analysis
09:50 - 10:15		Coffee Break
10:15 - 11:35		**Session 9 - Competing risks and multistate models** Chair: Pedro Miranda Afonso
	Salvatore Battaglia Palermo, Italy	Extending the vertical model: An alternative approach to competing risks with clustered data
	Sam Doerken Freiburg, Germany	Patient disposition in clinical trials: Addressing competing risks with stacked probability and proportion plots
	Marta Spreafico Leiden, The Netherlands	Discrimination performance in illness-death models with interval-censored disease data
	Yujun Xu Munich, Germany	Transitions, sojourns, and bias: Simulation insights for transplant strategies in leukemia
11:35 - 11:45		Short Break
11:45 - 12:45		**Keynote** Chair: Kathrin Möllenhoff
	Dennis Dobler Dortmund, Germany	Resampling options in survival and event history analysis
12:45 - 13:00		Closing Remarks, Best Talk and Poster Award

Conference information

Public transport to and in Bonn

Bonn is easily accessible by both air and rail. Travelers can arrive by plane via Cologne Bonn Airport (Köln-Bonn Airport, CGN) or by train at Bonn Central Station, which is well connected to major cities. If you arrive by plane, you can conveniently take the Airport-Express-Bus SB60, which will take you to Bonn Central Station in approximately 30 minutes. For departure times of the Airport-Express-Bus SB60 and train connections, we refer to the website of the Deutsche Bahn (DB) or the free DB navigator app (downloadable via the QR Codes below).

Download on the App Store

GET IT ON Google Play

At the registration desk, you will receive a public transport ticket for use within Bonn (provided that you indicated this in ConfTool during registration). This ticket will be valid during the conference dates (19th - 21st March 2025). It allows you to travel on all public transport services within the city of Bonn. Please note that the ticket is only valid within the city limits of Bonn. We refer to the **white-spaced area** on this map for further details. Please be aware that the Airport-Express-Bus SB60 is not included in the ticket. The valid travel zone covers the Godesburg restaurant and the Deutsches Museum, where social activities will take place.

Conference venue

The 10th Survival Analysis for Junior Researchers Conference will take place at the University Club Bonn:

University Club Bonn e.V.
Konviktstr. 9
53113 Bonn
Germany
https://www.uniclub-bonn.de

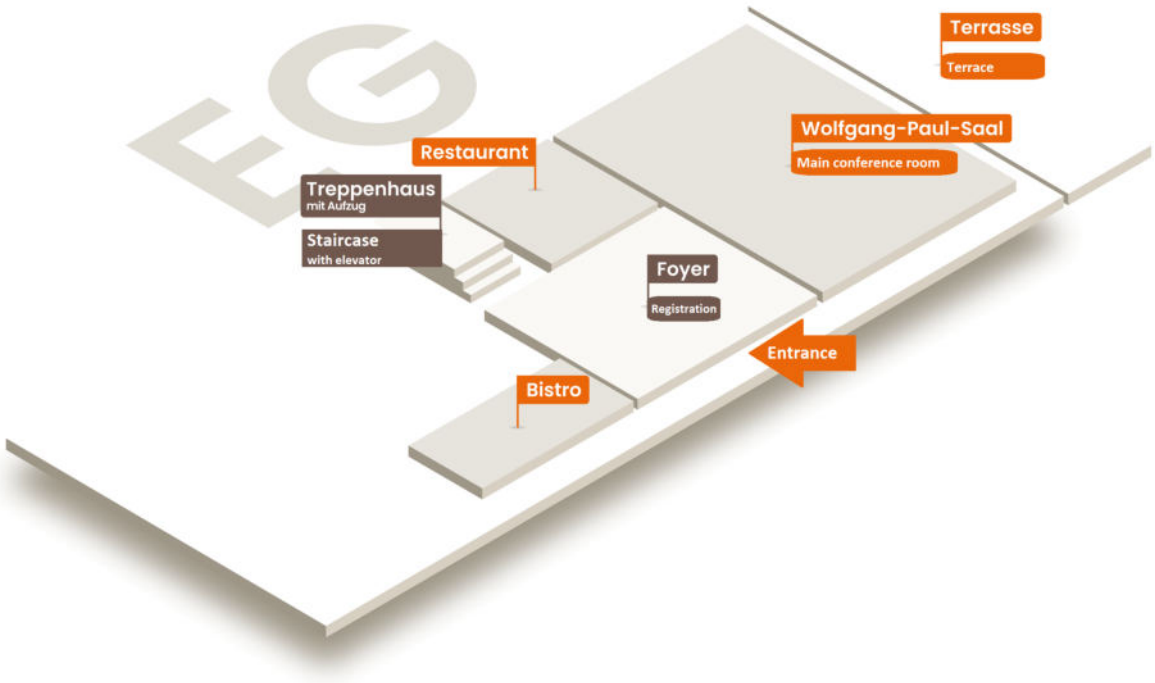
Universitätsclub Bonn

The University Club Bonn is a versatile venue close to the banks of the River Rhine. Its central location provides easy access to the city's cultural attractions and entertainment options, making it an ideal

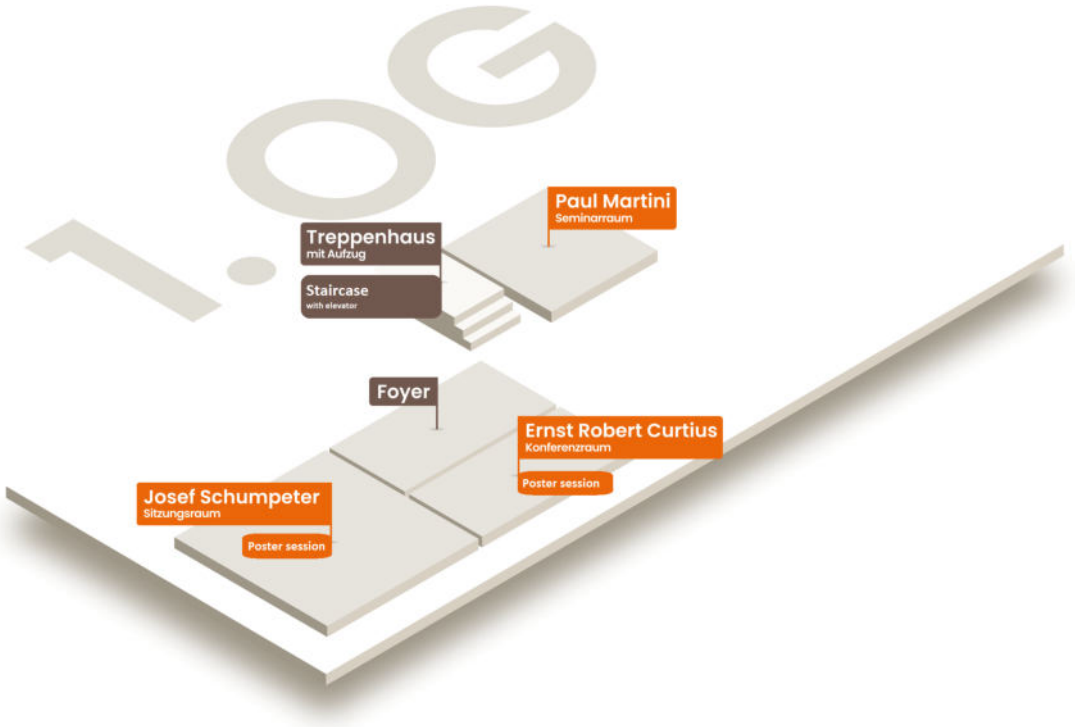
starting point for exploring Bonn's rich heritage and lively atmosphere. The closest tram station is *Universität / Markt*. Please scan the QR code above to view the venue on Google Maps and get directions for easy navigation. We recommend using the DB navigator app for bus, train and tram departure times (please see QR code on page 6).

The conference's main session will be held in the Wolfgang-Paul-Saal on the ground floor ("EG"). The room is situated to the right of the main entrance, as indicated on the map below.

Coffee, other drinks and snacks will be served in each break which is indicated as coffee break in the conference program (see pages 2 to 5). In the short breaks, coffee and other drinks will be available. The breaks will take place at the Foyer (see map below). Lunch will be available in the Restaurant on the ground floor.



The Poster session will take place in the Josef Schumpeter room and the Ernst Robert Curtius room on the first floor ("1. OG") which can be found on the map below.



The conference fee includes drinks and snacks during coffee and short breaks, as well as lunch on Wednesday and Thursday. A fingerfood buffet and drinks will be available during the poster session on Wednesday, also covered by the conference fee. On Friday, we will provide you with a complementary lunch package for your way home.

During the conference, you are welcome to use the Ernst Robert Curtius room on the first floor ("1. OG", see map above) for meetings or work.

The University Club offers free Wi-Fi, accessible either through eduroam (using your usual username and password) or via the gast-bonnet Wi-Fi. To connect to the gast-bonnet Wi-Fi, please use your personal username and password that you can find in your conference bag.

Prizes for best oral and poster presentations

Springer has generously provided the prizes for the best poster and best oral presentations. Winners will be awarded a free textbook to be selected from the Springer Series in Statistics or Springer Texts in Statistics. Further details will be provided to the winners after the award session.

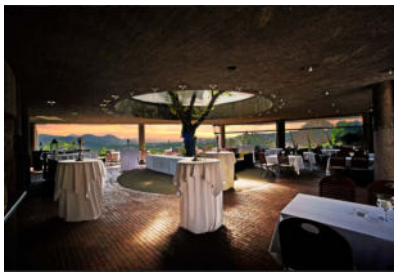
Artificial intelligence exhibition

The Deutsches Museum Bonn is currently hosting an exhibition on artificial intelligence ("Mission AI"), and we have reserved exclusive slots for our group on March 19, 2025 at 15:30. We will gather at 14:30 in the Foyer to travel to the museum. Attendees who have selected to join the AI exhibition will find a Deutsches Museum icon on their conference badge. Please scan the QR code to the left to view the venue at Google Maps and get directions for easy navigation. We recommend using the DB navigator app for bus, train and tram departure times (please see QR code on page 6).

Conference dinner

The SAfJR 2025 conference dinner will take place at the historic restaurant at Godesburg Castle on March 20, 2025, at 19:00. We look forward to sharing this special occasion with all attendees. Your ticket for public transport can be used to travel to the Godesburg. The closest tram station is *Plittersdorfer Straße*. Alternatively, the nearest train station is *Bonn Bad Godesberg* and the nearest bus station is *Bonn Bad Godesberg Bahnhof / Rheinallee*.
Please scan the QR code to view the venue on Google Maps and get directions for easy navigation. We recommend using the DB navigator app for bus, train and tram departure times (please see QR code on page 6). Attendees who have booked the conference dinner will find a Godesburg icon on their conference badge. The conference dinner will be served in buffet style, with all dietary restrictions taken into account. Alcoholic and non-alcoholic drinks are included in the conference dinner fee from 19:00 to 22:00. You are welcome to stay at the Godesburg after 22:00, but please note that drinks will need to be paid by yourselves after that time.





Organizing team

Alina Schenk is a PhD candidate in medical biometry at the IMBIE. Her research addresses predictive modeling of time-to-event outcomes using pseudo-values and machine learning techniques.

David Köhler is a PhD candidate in medical biometry at the IMBIE. His research interests include machine learning, time-to-event analysis, and regression modeling.

Jochen Kamuf is a member of the IT team of the IMBIE, which is responsible for overseeing the institute's IT infrastructure.

Prof. Dr. Matthias Schmid is the director of the IMBIE. His research interests include regression modeling, statistical learning, and time-to-event-analysis.

Tanja Hoffmann is responsible for management and coordination at the IMBIE.

Vanessa Basten is a research associate at the IMBIE. She works on pseudo-value regression and statistical modeling of observational data.

Keynote sessions

Regression analysis with jack-knife pseudo-observations

Morten Overgaard [1]

[1] Aarhus University, Denmark

Abstract:
More than twenty years ago, a pseudo-observation approach to regression in survival analysis was suggested. Such an approach allows for handling incomplete observation of a relevant outcome variable by transforming the available data into pseudo-observations that can replace the potentially unobserved outcomes in a regression analysis. The pseudo-observations are here the jack-knife leave-one-out pseudo-values from a suitable estimator.

The original motivation of the approach was in a multi-state setting with right censoring. It has since been suggested for settings with complications such as left-truncation, interval censoring or recurrent events.

We will review this sort of approach and attempt to find answers to questions such as: How do I use it? When does it work? Why is left-truncated data a challenge? Is variance estimation an issue? How does it compare to similar approaches? Where is it going from here? Are pseudo-observations fun?

Morten Overgaard is an Assistant Professor in the Department of Public Health at Aarhus University, where he has been contributing to research and education since 2020. He earned his PhD in 2018 from Aarhus University's Graduate School of Health, following an MSc in Statistics and a BSc in Mathematics, both from Aarhus University. Morten's academic career includes roles as a postdoctoral researcher, PhD student, and research assistant in the Department of Public Health, as well as experience as a computing assistant and instructor in the Department of Mathematical Sciences. His expertise spans statistics and public health research, with a strong foundation in mathematics and data analysis.

Causal prediction of time-to-event outcomes

Nan van Geloven[1]

[1] Leiden University, The Netherlands

Abstract:
A key aim of clinical prediction models (or prognostic models) is to provide individualized risk estimates that support patients and doctors in making treatment decisions. However, as most prediction models are derived from observational data where some individuals already received the treatment the model aims to inform, standard prediction methods often fail to provide the necessary information for this.

Causal predictions (also called counterfactual predictions or predictions under intervention) have recently gain traction as an alternative way to support treatment decisions. These are estimates of risks under specified treatment strategies, for example a patient's baseline risk 'if they do not initiate the treatment' or the risk 'if they do initiate the treatment'. A major challenge in estimating and evaluating causal predictions in observational data is confounding adjustment.

In this talk, I will outline how causal inference methods such as marginal structural models and the clone-censor-reweight approach can be adapted for the purpose of developing and evaluating causal predictions of time-to-event outcomes.

Nan van Geloven is an Assistant Professor at Leiden University Medical Center with expertise in biostatistics, applied mathematics, and is also recognized for her extensive work in the field of causality. She has experience in statistical consultancy, research, and education, including developing online teaching tools like Wiki Biostatistics and Practical Biostatistics. Previously, she held roles at the Academic Medical Center in Amsterdam and Gupta Strategists. Nan holds a PhD in Biostatistics from the University of Amsterdam and a MSc in Applied Mathematics from Delft University of Technology.

Resampling options in survival and event history analysis

Dennis Dobler[1]

[1] Dortmund University, Germany

Abstract:
Resampling plays a versatile role in every branch of statistics. For example, it is used to compute variance estimates, to avoid the suboptimal normal distribution approximation in inference methods, or as bagging (bootstrap aggregating) in random forest algorithms. There exists an abundance of resampling methods, in particular, the classical, weighted, multiplier, and parametric bootstraps, and random permutation.

In survival analysis, the issue of incomplete data demands special attention when it comes to resampling. For instance, Akritas (1986, JASA) found that Efron's suggestion (1981, JASA) to draw with replacement from the censored data points before recomputing the Kaplan-Meier estimator allows the construction of asymptotically valid confidence bands for the survival function. And also that, in contrast, Reid's approach (1981, Biometrika) to independently resample from the Kaplan-Meier curve alters the covariance structure of the resampled Kaplan-Meier estimator; hence, it should not be used for constructing such confidence bands.

This talk will provide an incomplete overview of various resampling procedures in different survival analytic applications, address some requirements and intuitions behind these procedures, and also discuss some computational aspects.

Photo copyright: Hesham Elsherif / TU Dortmund

Dennis Dobler is a Junior Professor and researcher in mathematical statistics and trustworthy data science, affiliated with the Faculty of Statistics and the Research Center for Trustworthy Data Science and Security at the University Alliance Ruhr. He holds a PhD in Mathematics from the University of Ulm and dual BSc and MSc degrees in Mathematics and Computer Science from Heinrich Heine University Düsseldorf. Previously, Dennis was an Assistant Professor of Mathematical Statistics at Vrije Universiteit Amsterdam, a postdoctoral researcher at the University of Ulm, and a visiting researcher at the University of Copenhagen. His research focuses on resampling methods, survival analysis, multivariate analysis, and asymptotic statistics.

Risk prediction using case-cohort samples: A scoping review and empirical comparison

Y. Li[1], R. Keogh[2], C. Kartsonaki[1]

[1] University of Oxford, United Kingdom
[2] Department of Medical Statistics, London School of Hygiene & Tropical Medicine, United Kingdom

E-mail: reub0433@ox.ac.uk

Abstract:
Risk prediction models are commonly developed using data from observational cohorts. When measuring certain variables across an entire cohort is infeasible due to high costs, using a case-cohort sample offers an alternative. The case-cohort design limits the full measurement of covariates to a randomly sampled subcohort along with all remaining cases, which offers an efficient strategy that maintains reasonable statistical power. However, developing risk prediction models in case-cohort data poses challenges due to its outcome-dependent sampling design, where the overrepresentation of cases can distort predictor-outcome relationships if not appropriately accounted for.

Most existing case-cohort analyses have used traditional weighted Cox regression models (Prentice, Self-Prentice, and Barlow), although several methods have been proposed in recent years to develop risk prediction models in case-cohort design, including multiple imputation techniques to leverage full cohort information (Keogh et al., 2013), inverse probability-weighted kernel machine methods (Payne et al., 2016), and Bayesian frameworks for case-cohort Cox regression (Yiu et al., 2021). While these new methods are promising, their practical applicability, advantages, limitations, and real-world performance remain unvalidated.

This study aims to bridge this gap by: (1) reviewing methodological advancements in case-cohort analysis, including variable selection, missing data imputation, sampling design, and model evaluation; (2) summarizing research that has developed risk prediction models with real-world case-cohort data; and (3) comparing available tools (e.g., R, Stata, and Python packages) for case-cohort analysis. Additionally, we plan an empirical comparison of these methods and software tools on case-cohort data from the China Kadoorie Biobank (CKB), allowing us to assess their practical strengths and limitations. Finally, we will discuss current limitations in the field and highlight areas for further development.

Keywords:
Risk prediction; Case-cohort; Model evaluation; Novel predictive biomarker.

Implication of the choice of time scales in survival analysis

J. Vilsmeier[1], G. Büchele[2], M. Rehm[2], D. Rothenbacher[2], J. Beyersmann[1]

[1] Institute of Statistics, Ulm University, Ulm
[2] Institute of Epidemiology and Medical Biometry, Ulm University, Ulm

E-mail: judith.vilsmeier@uni-ulm.de

Abstract:
The choice of time scale is an important and often discussed topic in time-to-event analysis. While the rule of thumb is to choose the natural time scale for the underlying problem, it is often not clear what this natural time scale is. This potentially leads to confusion and if researchers pick the ”wrong” time scale for the underlying problem, this can lead to biased results. Two often discussed time scales are time since study entry and time since birth, i.e. age, but there are also other possible time scales. The time scale we focus on during this talk is the so-called calendar time. Its origin is an arbitrary day before the study entry of the first patients, and it is useful if calendar time holds important information in addition to just the time from study entry. However, we will illustrate situations where it can lead to an inflated estimation of the hazard ratio and overoptimistic p-values when using the Cox proportional hazard model. In this talk a real data example from the EvaCoM project is used to demonstrate this hazard ratio inflation and the reasons why it occurs. In this project, the impact of quality audit of healthcare providers was of interest, possible time scales being calendar time in which the audit acts or a patient’s time-to-event. It serves as an example to raise awareness to why researchers should be careful when deciding which time scale to use, by highlighting a situation in which the choice of time scale might not be clear but impactful. Additionally, possible solutions for scenarios where more than one time scale seem to be important are proposed.

Keywords:
Time scales; Real-time approach; Cox proportional hazards model; Poisson regression; Matching.

Leveraging cancer incidence for lead time estimation in cancer screening programmes

B. Vratanar[1], M. Pohar Perme[1]

[1] Institute for Biostatistics and Medical Informatics, Slovenia

E-mail: bor.vratanar@gmail.com

Abstract:
In cancer screening programmes, participants are regularly screened every few years using blood tests, urine tests, or medical imaging to detect cancer at an earlier time, when it is presumed to be more curable. Without screening, cancer would likely progress undetected until symptoms appear. The interval between early detection and the eventual onset of symptoms, had screening not been conducted, is known as lead time. Due to lead time, the survival time for screen-detected cancers is artificially extended compared to the cancers detected based on symptoms as they are being followed from an earlier point in time, resulting in a biased comparison. Understanding and estimating lead time is thus crucial for researchers to mitigate lead time bias in cancer screening studies.

Estimating lead time is challenging because it is a hypothetical random variable that can only be inferred indirectly. In our study, we introduce a novel method for estimating lead time, using a data source previously untapped for this purpose—cancer incidence. We hypothesize that earlier detection of cancer due to screening should result in observable shift in cancer incidence rates, stratified by age and year of diagnosis. Our method leverages this information, and estimates lead time using a maximum likelihood estimator. In principle, the user specifies the distribution of lead time, and the method finds the parameters that best fit the observed shift in cancer incidence.

Our approach is flexible, allowing for the inclusion of additional covariates and accounting for overdiagnosis. The data required for this method are routinely available from cancer registries and provided by population tables, making it easy for implementation. We validated our method through simulations and applied it to data from the Slovenian breast cancer screening programme, demonstrating its effectiveness and utility.

Keywords:
Cancer screening programmes; Lead time; Cancer incidence.

References:
B. Vratanar and M. Pohar Perme (2023): Evaluating cancer screening programs using survival analysis. Biometrical Journal, 65(7), 2200344.

Capturing subgroup-specific time-variation in covariate effects in Cox-type hazard regression models

N. Hagemann[1,2], T. Kneib[3], K. Möllenhoff[1,2]

[1] Institute of Medical Statistics and Computational Biology, Faculty of Medicine, University of Cologne, Germany
[2] Division of Mathematics, Department of Mathematics and Computer Science, University of Cologne, Germany
[3] Chair of Statistics, Georg-August-University Göttingen, Germany

E-mail: niklas.hagemann@uni-koeln.de

Abstract:
One of the major topics in survival analysis is analyzing the effect covariates have on the survival time. Frequently, these covariate effects are observed to be either time-varying, heterogeneous, i.e. patient, treatment or subgroup specific, or even both. While the standard model, the Cox proportional hazards model, allows neither time-varying nor heterogeneous effects, several extensions to the Cox model as well as alternative modeling frameworks have been introduced. However, none of these studies includes heterogeneously time-varying effects of covariates. Such effects occur if a covariate influences the survival time not only in a heterogeneous and time-varying manner, but this time-variation is heterogeneous, too.

In this talk we propose to model such effects by introducing heterogeneously time-varying coefficients to piece-wise exponential additive mixed models. We deploy functional random effects, also known as factor smooths, to model such coefficients as the interaction effect of heterogeneity and time-variation. Our approach allows for non-linear time-effects due to being based on penalized splines and uses an efficient random effects basis to model the heterogeneity. Using a penalized basis prevents overfitting in case of absence of such effects. In addition, the penalization mostly solves the problem of choosing the number of intervals which is usually present in piece-wise exponential approaches. The practical relevance is outlined by presenting a brain tumor case study. Finally, we demonstrate the superiority of our approach in comparison to competitors by means of a simulation study.

Keywords:
Hazard regression; Piece-wise exponential models; Functional random effects; Heterogeneity; Time-varying covariate effects.

Dynamic prediction of survival benefit to inform liver transplant decisions in hepatocellular carcinoma

P. Miranda Afonso[1], H. Liu[2], M. Molinari[3], D. Rizopoulos[1]

[1] Department of Biostatistics, Erasmus Medical Center, Netherlands
[2] Starzl Transplant Institute, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania, US
[3] J.C. Walter Jr Transplant Center, Houston Methodist Hospital, Houston, Texas, US

E-mail: p.mirandaafonso@erasmusmc.nl

Abstract:
Liver transplantation provides the best survival outcomes for patients with early hepatocellular carcinoma (HCC). However, as the demand far exceeds the number of available organs, it is crucial to identify the patients who will benefit most from liver transplantation. The α-fetoprotein (AFP) level, the radiographic tumour burden score (TBS), and the model for end-stage liver disease (MELD) score are routinely measured to monitor HCC progression and liver dysfunction to inform transplantation decisions. The transplant-related survival benefit has been proposed as a comprehensive metric combining post-transplant and waiting list life expectancies. Despite recent advances, existing methods of estimating this metric disregard the full longitudinal information, the observational nature of the data, and the influence of time-varying confounders.

Our primary goal is to dynamically predict the individualized transplant-related survival benefit in HCC patients to improve transplant prioritization. We analyse data from 7,471 waitlisted and 4,786 transplanted HCC patients listed in the US Scientific Registry for Transplant Recipients (SRTR) between March 2002 and March 2022.

We propose a Bayesian joint model that associates the risk of death, before and after transplantation, with the pre-transplant AFP level, TBS, and MELD score trajectories. We establish the necessary assumptions to unbiasedly estimate the causal transplantation effects from the observational data at hand. Using the postulated joint model we predict the pre- and post-transplant five-year survival probabilities for each patient at risk at a given time. From these, we derive the posterior distribution of the causal transplantation benefit. This distribution captures the full range of potential benefits, helping to identify patients most likely to derive the greatest benefit from an available organ. The individual benefits can be dynamically updated as new marker measurements become available. The model is implemented in the R package JMbayes2.

Our results demonstrate that AFP, TBS, and MELD each have distinct associations with the risk of death before and after transplantation, confirming the value of the new approach. Despite the observational nature of the SRTR data, our model can unbiasedly estimate the causal effect of transplantation on survival, leading to unbiased individualized transplant-related benefit distributions.

Our modelling strategy ensures a fair allocation of the limited number of liver transplants, optimizing the use of available organs and improving the overall survival of all waitlisted patients.

Keywords:
Dynamic prediction; Joint model; Observational data; Survival benefit; Survival analysis.

Dynamic prediction with numerous longitudinal predictors: How to combine the best of both worlds (landmarking and joint modelling) through penalized regression calibration

M. Signorelli[1]

[1] Leiden University, The Netherlands

E-mail: msignorelli.papers@gmail.com

Abstract:
Longitudinal and high-dimensional data are nowadays common in biomedical research. Repeated measurements data carry important information about ageing and disease progression, and this information can be used to dynamically update predictions of survival outcomes, such as the onset of dementia, cancer relapses, and death. Traditional approaches to dynamic prediction include joint modelling, which becomes computationally burdensome as the number of longitudinal predictors increases, and LOCF landmarking, which is computationally straightforward but only uses data from the last available observation.

In this talk I will introduce penalized regression calibration (PRC) (Signorelli et al., 2021; Signorelli, 2024), a new statistical method that strives to strike a balance between the mathematical elegance of joint models, and the simplicity of landmarking.

In short, PRC specifies a conditional model for the probability of experiencing an event in the future, given that a subject has not experienced the event up until a given landmark time l. PRC allows flexible modelling of the longitudinal data gathered up to the landmark through mixed-effects models, and it models the conditional survival function $S(t|l)$ using appropriate summaries of the longitudinal covariates as predictors (alongside relevant baseline / time-constant covariates) in a Cox model. Estimation of the Cox model can be performed using penalized likelihood: this allows to reduce the risk of overfitting the available data, and makes it possible to handle a large number of predictors, both in low- and high-dimensional settings.

After illustrating how PRC works, I will show how it can be easily implemented using the R package pencal (Signorelli, 2024), and discuss how to properly validate the predicted performance of the fitted models. I will present the results of applications that show that PRC can achieve a predictive performance that is comparable to that of joint models, but in a much more computationally-efficient way. I will conclude by presenting our ongoing work to enhance the flexibility of PRC so that it can better handle discrete longitudinal predictors, interval censoring, and competing risks.

Keywords:
Dynamic prediction; Longitudinal data; Mixed-effects models; Penalized likelihood; Risk prediction models.

References:

M. Signorelli, P. Spitali, C. A. Szigyarto, The MARK-MD Consortium and R. Tsonaka (2021): Penalized regression calibration: A method for the prediction of survival outcomes using complex longitudinal and high-dimensional data. Statistics in Medicine, 40(27), 6178–6196.

M. Signorelli (accepted, 2024): pencal: An R Package for the Dynamic Prediction of Survival with Many Longitudinal Predictors. To appear in: The R Journal, Preprint: arXiv:2309.15600.

Bootstrap-based inference for pseudo-value regression models

S. Mack[1,2,3], D. Dobler[1,3], M. Overgaard[4]

[1] TU Dortmund University, Germany
[2] Otto von Guericke University Magdeburg, Germany
[3] Research Center Trustworthy Data Science and Security, Germany
[4] Department of Public Health - Department of Biostatistics, Aarhus University, Denmark

E-mail: simon.mack@tu-dortmund.de

Abstract:
Generalized estimating equations (GEE) are a popular method to model the effects of covariates on various estimands, which only rely on the specification of a functional relationship without the need of restrictive distributional assumptions. However, if the response variable is not fully observable, e.g. in the case of time-to-event data, the GEE approach is not directly applicable.

Andersen et al. (2003) proposed to replace the partially unobservable response variables by jackknife pseudo-observations, and Overgaard et al. (2017) showed that the resulting parameter estimates are consistent and asymptotically normal under very general conditions. For further inference about the parameter vector an estimator of the asymptotic covariance matrix is necessary. But due to the dependence of the pseudo-observations, the limiting covariance matrix is highly complicated and the usual sandwich estimator seems to be inconsistent (Jacobsen and Martinussen, 2016; Overgaard et al., 2018). Overgaard et al. (2017) proposed an alternative estimator which incorporates the dependence of the pseudo-observations and performs well in medium to large samples. These results would in principle allow for the construction of tests for general linear hypotheses about the parameters. However, mainly confidence intervals for individual parameters or simple contrasts, e.g. risk differences, have been considered in the past. In this talk we aim to bridge this gap by introducing different test statistics for general linear hypotheses in pseudo-value regression models.

To improve the small sample performance of these tests we discuss different bootstrap methods for pseudo-observations as well as possible extensions to multiple testing problems and simultaneous confidence intervals for contrasts.

Acknowledgements:
We would like to thank Marc Ditzhaus for his invaluable collaboration and guidance in the early phase of this work. Sadly, he has deceased and he could not complete this work together with us.

Keywords:
Generalized estimating equation; Jack-knife pseudo-values; Time-to-event analysis; Bootstrap; Multiple testing.

References:

P. K. Andersen, J. P. Klein and S. Rosthøj (2003): Generalised linear models for correlated pseudo-observations, with applications to multi-state models. Biometrika, 90.1, 15-27.

M. Overgaard, E. T. Parner and J. Pedersen (2017): Asymptotic theory of generalized estimating equations based on jack-knife pseudo-observations. The Annals of Statistics, 45.5, 1988-2015.

M. Jacobsen and T. Martinussen (2016): A note on the large sample properties of estimators based on generalized linear models for correlated pseudo-observations. Scandinavian Journal of Statistics, 43.3, 845-862.

M. Overgaard, E. T. Parner and J. Pedersen (2018): Estimating the variance in a pseudo-observation scheme with competing risks. Scandinavian Journal of Statistics, 45.4, 923-940.

Implications of pseudo-observations in prognostic modelling: Addressing left truncation

N. Murunga[1], S. Booth[1], M. Rutherford[1]

[1] Biostatistics Research Group, Department of Population Health Sciences, University of Leicester

E-mail: nm545@leicester.ac.uk

Abstract:

Background: In many time-to-event prognostic models, the Cox model is used, commonly assuming constant covariate effects over the full follow-up period (McLernon et al., 2023). However, when interest is in fact in estimating prognosis at specific future timepoints after diagnosis, then Pseudo-Observations (POs), introduced by Andersen et al. (2010), can be used to directly model covariate effects on survival at the timepoint of interest (Andersen et al., 2010).

We aim to utilize POs in prognostic modelling for specified time points when the objective is to provide an up-to-date estimate of long-term survival, which is typically solved with period analysis. In this approach, only risk times and events within a defined recent period window inform survival estimates, providing more up-to-date predictions than standard full-cohort approaches, which may underestimate survival as prognosis improve over time (Booth et al., 2023).

However, defining this period window introduces left-truncated data, as individuals enter the study after a specified calendar time point. This complicates analysis because late-entry individuals receive POs without contributing to the Kaplan-Meier (K-M) estimate at the time of interest, potentially distorting survival probabilities (Grand et al., 2019).

This study aims to refine POs for better alignment with K-M estimates in left-truncated scenarios, enhancing prognostic accuracy and fully incorporating both right-censored and left-truncated data, resulting in updated survival estimates within period window analysis.

Methods: We demonstrate this approach with a prognostic model using historical data for colon cancer patients aged 18-99, diagnosed between 1975 and 1994. A stratified approach to POs is introduced to address left truncation in prognostic modelling. First, POs are calculated without delayed entry to build a model categorizing individuals into risk groups based on covariates. A period window then introduces delayed entry, and POs are recalculated within each risk group, averaged, and compared with K-M estimates for agreement. Once consistency is established, the new POs are used to update the baseline of the prognostic model in period analysis. We compare a range of choices for how many risk groups are necessary to achieve good agreement.

Results and Discussion: This approach demonstrates that average PO estimates within risk groups align closely with K-M estimates, supporting the use of stratified POs in prognostic models with period window analysis and left truncation. Updated POs can refine the baseline of prognostic models, helping account for survival improvements over time. However, a more general method, independent of predefined risk groups or model, remains necessary. Alternative methods like inverse probability of censoring weighting, offer an alternative approach.

Keywords:
Pseudo-observations; Period window; Left truncation; Prognostic modelling; Delayed entry.

References:

D. J. McLernon, D. Giardiello, B. Van Calster, L. Wynants, N. van Geloven, M. van Smeden et al. (2023): Assessing Performance and Clinical Usefulness in Prediction Models With Survival Outcomes: Practical Guidance for Cox Proportional Hazards Models. Ann Intern Med., 176(1):105–14.

P. K. Andersen, M. P. Perme (2010): Pseudo-observations in survival analysis. Stat Methods Med Res., 19(1):71–99.

S. Booth, S. I. Mozumder, L. Archer, J. Ensor, R. D. Riley, P. C. Lambert et al (2023): Using temporal recalibration to improve the calibration of risk prediction models in competing risk settings when there are trends in survival over time. Stat Med., 42(27):5007–24.

M. K. Grand, H. Putter, A. Allignol,P. K. Andersen (2019): A note on pseudo-observations and left-truncation. Biom J Biom Z., 61(2):290–8.

Estimation and variables selection in a joint model of survival times and longitudinal data with random effects

A. Caillebotte[1,2], E. Kuhn[1,2], S. Lemler[1,3]

[1] Université Paris-Saclay
[2] INRAE UR MaIAGE
[3] CentraleSupélec, Laboratoire MICS

E-mail: antoine.caillebotte@inrae.fr

Abstract:
This work considers a joint survival and mixed-effects model to explain the survival time from longitudinal data and high-dimensional covariates. The longitudinal data is modeled using a nonlinear mixed effects model, whose regression function serves as a link function incorporated into a Cox model as a covariate. In that way, the longitudinal data is related to the survival time. Additionally, the Cox model takes into account high-dimensional covariates. The main objectives of this research are two-fold: identify the relevant covariates that contribute to explaining survival time and estimate all unknown parameters of the joint model, in particular the strength of the link between the two models. For that purpose, we consider the estimate obtained by maximizing a LASSO penalized likelihood. To tackle the optimization problem, we implement a pre-conditioned stochastic gradient to handle the latent variables of the nonlinear mixed-effects model associated with a proximal operator to manage the non-differentiability of the LASSO penalty. Variable selection lies on the LASSO penalization and therefore on a regularization parameter chosen according to the eBIC criterion. The latter is better suited to the high-dimensional context. We provide a simulation study showcasing the performance of the proposed variable selection and parameter estimation method.

Keywords:
High dimension; Variable selection; Computational statistics; Joint model; Mixed effects.

Sign-flip test for coefficients in the Cox regression model

R. De Santis[1], J. J. Goeman[2], H. Putter[2], L. Finos[3]

[1] University of Siena, Italy
[2] Leiden University Medical Center, The Netherlands
[3] University of Padova, Italy

E-mail: riccardo.desantis2@unisi.it

Abstract:
Cox regression model is a popular tool in survival analysis, whose aim is to quantify the impact of covariates on the survival times. The relevance of the coefficients is usually tested through a parametric test. However, the properties of this test are only asymptotical and can show a slow convergence to the nominal level. We propose a different approach to perform the test on the coefficients based on sign-flipping of the score contributions. Simulations show a faster convergence to the nominal level of the test based on the proposed method. Further, we embed the new test in a permutation-based framework to tackle the case of multiple coefficients testing, which turns out to be relevant especially in high-dimensional problems. We show the potential of our proposal through a real data application in genomics.

Keywords:
Sign-flip; Semiparametric test; Cox model; Multiple testing.

References:

R. De Santis, J. J. Goeman, J. Hemerik, S. Davenport and L. Finos (2022): Inference in generalized linear models with robustness to misspecified variances. arXiv preprint arXiv:2209.13918, (on revision).

R. De Santis, J. J. Goeman, S. Davenport, J. Hemerik and L. Finos (2024): Permutation-based multiple testing when fitting many generalized linear models. arXiv preprint arXiv:2403.02065, (on revision).

Targeted learning with right-censored data using the state learner

A. Munch[1], T. Gerds[1]

[1] University of Copenhagen, Denmark

E-mail: a.munch@sund.ku.dk

Abstract:
Targeted or debiased machine learning provides a methodology for combining data-adaptive estimators with asymptotically valid inference for interpretable estimands. In particular, using a super learner as a data-adaptive model selector offers a general framework for obtaining valid statistical inference without relying on a single pre-specified model. Super learning evaluates model performance using cross-validation. However, cross-validation based on right-censored data typically depends on a pre-specified model for the censoring distribution, which can be challenging to provide, especially for observational data. To address this, we propose a new super learner, the state learner, which jointly evaluates the performance of models for both the outcome and censoring distributions. The state learner uses the data to select a pair of models that are optimal for predicting the state-occupation probabilities characterizing the observed data distribution. This approach readily extends to settings with competing risks and is particularly well suited for use in combination with targeted learning. We discuss the theoretical properties of the state learner and demonstrate how it can be integrated with targeted learning for estimation of low-dimensional, interpretable estimands in a competing risks model observed under right-censoring.

Deep learning for survival analysis: A review

S. Wiegrebe[1]

[1] LMU Munich, Germany

E-mail: simon.wiegrebe@stat.uni-muenchen.de

Abstract:
The influx of deep learning (DL) techniques into the field of survival analysis in recent years has led to substantial methodological progress; for instance, learning from unstructured or high-dimensional data such as images, text or omics data.

We conducted a comprehensive systematic review of 61 DL-based methods for time-to-event analysis, by defining dimensions according to which we classified all methods. These dimensions include survival-related aspects like the ability to handle different survival tasks (different types of censoring and truncation schemes, competing risks, multi-state models, etc.); handling of non-proportional hazards (time-varying effects and features); neural network-related aspects like architecture; as well as estimation-related aspects (model family, target of estimation, etc.).

On the one hand, this work provides a snapshot of the current state of the field and identifies potential gaps for future research, on the other hand it provides a scheme according to which future methods can be categorized.

From a technical perspective, we found that most methodologically innovative methods were survival-specific applications of novel methods developed in other areas of DL, such as computer vision or NLP. This usually yielded a more flexible estimation of associations of (structured and unstructured) features with the outcome, rather than solving survival-specific problems. Outcome types beyond right-censoring and competing risks, as well as handling of missing values, were rarely addressed and only little attention has been paid to optimization (e.g., choice of optimizers, tuning of hyperparameters, or neural architecture search). Additionally, there were some challenges which are specifically DL-related, in particular batching.

From an application-centered perspective, DL-based survival methods have mostly been deployed in estimating patient survival based on medical images or (multi-)omics data, some methods being explicitly motivated by a specific clinical use case (e.g., a specific cancer type). Other areas of application of DL-based survival methods included improved estimation of prognostic indices and of recurrence after cancer surgery.

In more general terms, we observed that the lack of openly accessible, high-dimensional, potentially multimodal datasets remains a major challenge for the development and training of novel DL-based survival methods.

In summary, deep survival methodology has advanced substantially in recent years and will certainly continue to benefit from developments in general DL, with big methodological advances being likely to swap over. The results of our review are summarized in an interactive online table (https://survival-org.github.io/DL4Survival), which can be extended through pull requests.

Keywords:
Deep learning; Systematic review; Comparison.

References:

S. Wiegrebe, P. Kopper, R. Sonabend, B. Bischl and A. Bender (2024): Deep learning for survival analysis: a review. Artificial Intelligence Review, 57(3), 65.

P. Kopper, S. Wiegrebe, B. Bischl, A. Bender and D. Rügamer (2022): DeepPAMM: deep piecewise exponential additive mixed models for complex hazard structures in survival analysis. In Pacific-Asia Conference on Knowledge Discovery and Data Mining (pp. 249-261). Cham: Springer International Publishing.

Copula based dependent censoring in cure models with covariates

M. Delhelle[1], A. El Ghouch[1], I. Van Keilegom[1,2]

[1] Institute of Statistics, Biostatistics and Actuarial Sciences, UCLouvain, Belgium
[2] Operations Research and Statistics Research Group, KU Leuven, Belgium

E-mail: morine.delhelle@uclouvain.be

Abstract:
In the field of survival data analysis, datasets that include both a cure fraction (i.e., individuals who will never experience the event of interest) and dependent censoring (loss to follow-up for a reason related to the event of interest before the occurrence of that event) are not scarce. It is therefore essential to consider appropriate models and methods in order to avoid biased estimators of the survival function or incorrect conclusions in clinical trials. Delhelle and Van Keilegom (2024) proposed a fully parametric mixture cure model for the bivariate distribution of survival and censoring times (T,C), which deals with all these features. The model depends on a parametric copula and on parametric marginal distributions for T and C. A significant advantage of this approach in comparison to existing approaches in the literature is that the copula which models the dependence between T and C is not assumed to be known, nor is the association parameter. Furthermore, the model allows for the identification and estimation of the cure fraction and the association between T and C, despite the fact that only the smallest of these variables is observable.

This talk presents an improvement of this model. Administrative censoring is considered separately from dependent censoring, and covariates are included in the model.

Keywords:
Copulas; Cure models; Dependent censoring; Covariates; Identifiability.

References:
M. Delhelle and I. Van Keilegom (2024): Copula based dependent censoring in cure models. TEST, (under revision), arXiv preprint arXiv:2403.07963.

Testing the effect of multiple covariates on cure rates in mixture cure models based on distance correlation

B. E. Monroy-Castillo[1], A. Jácome[1], R. Cao[1]

[1] Universidade da Coruña, Spain

E-mail: b.mcastillo@udc.es

Abstract:
In survival analysis, certain scenarios involve cases where not all individuals are at risk of experiencing the event of interest. Advanced methods have been developed to address such data, commonly referred to as cure model analysis. A cure model allows for direct modeling of the cure proportion and the effect of covariates on it. Based on how the cure proportion is introduced, cure models can be broadly categorized into two types: mixture cure models and non-mixture cure models. Mixture cure models enable the estimation of both the probability of being cured and the survival function for uncured individuals. A key objective in these models is to determine whether covariates influence the cure rate.

To develop a new nonparametric test, the focus shifts to a novel measure: distance correlation. Distance correlation has several important properties, including its applicability to vectors X and Y in arbitrary dimensions, and the fact that a distance correlation of zero characterizes independence. One notable extension of distance correlation is martingale difference correlation, which evaluates deviations from conditional mean independence between a scalar response variable Y and a vector predictor variable X. Moreover, martingale difference correlation and its empirical counterpart retain many advantageous properties of distance correlation and sample distance correlation.

This study proposes a new test for assessing the significance of multiple covariates, leveraging martingale difference correlation. The effectiveness of the proposed test is evaluated through a Monte Carlo simulation study under various scenarios, and the method is applied to a rheumatoid arthritis dataset.

Keywords:
Martingale difference correlation; Permutation test; Censoring; Non-parametric test.

References:
B. E. Monroy-Castillo, A. Jácome and R. Cao (2024): Improved distance correlation estimation. Applied Intelligence. Advance online publication. arXiv preprint arXiv:2403.07963.

The sicure R package: Single-index mixture cure models

B. Piñeiro-Lamas[1], A. López-Cheda[1], R. Cao[1]

[1] Universidade da Coruña, Spain

E-mail: b.pineiro.lamas@udc.es

Abstract:
In survival analysis, there are situations in which not all subjects are susceptible to the final event. For example, if the event is a cancer therapy-related adverse effect, there will be a fraction of patients (considered as cured) that will never experience it. Mixture cure models address this by estimating both the probability of cure and the survival function for the uncured subjects. In the literature, nonparametric estimation of these functions focuses on continuous univariate covariates. However, in clinical practice, it is common to collect several patient characteristics and even medical images. The R package sicure provides a set of functions related to the implementation of single-index mixture cure models, that can handle a vector covariate under the assumption that the survival function depends on it through an unknown linear combination that can be estimated by maximum likelihood. This approach can be easily extended to functional covariates. The implementation of a nonparametric estimator for the density function of the uncured individuals is also included. Although the use of this package is illustrated with a medical dataset, it may be useful in any other field that involves a time variable, an uncensoring indicator, more than one covariate and the presence of a cure fraction.

Keywords:
Cardiotoxicity; Dimension reduction; Functional data; Kernel estimation.

References:

B. Piñeiro-Lamas, A. López-Cheda, R. Cao, L. Ramos-Alonso, G. González-Barbeito, C. Barbeito-Caamaño and A. Bouzas-Mosquera (2023): A cardiotoxicity dataset for breast cancer patients. Scientific Data, 10, 527, doi:10.1038/s41597-023-02419-1.

B. Piñeiro-Lamas, A. López-Cheda, R. Cao, L. Ramos-Alonso, G. González-Barbeito, C. Barbeito-Caamaño and A. Bouzas-Mosquera (2023): BC_cardiotox: A cardiotoxicity dataset for breast cancer patients. doi:10.6084/m9.figshare.22650748.

B. Piñeiro-Lamas, A. López-Cheda and R. Cao (Submitted): Maximum likelihood estimation in single-index mixture cure models.

Testing for sufficient follow-up in survival data with covariates

E. Musta[1], I. van Keilegom[2], T. P. Yuen[1]

[1] University of Amsterdam, The Netherlands
[2] KU Leuven, Belgium

E-mail: t.p.yuen@uva.nl

Abstract:
Survival data in the presence of a cure fraction has recently attracted growing interest from both methodological and application perspectives. To estimate the fraction of 'immune' or 'cured' subjects who will never experience the event of interest, it is necessary to have a sufficiently long follow-up period. A few statistical tests have been introduced to test the assumption of sufficient follow-up, indicating that the right extreme of the censoring distribution exceeds that of the survival time of the uncured subjects. A relaxed notion of 'practically' sufficient follow-up has been proposed recently, suggesting that the follow-up would be considered sufficiently long if the probability for the event happening after the end of the study is very small. However, all these tests do not consider covariate information, which might affect the cure rate and the survival times. We develop a novel statistical test for 'practically' sufficient follow-up that accounts for covariates. Our approach relies on the assumption that the density in the tail of the conditional distribution of uncured survival time is a non-increasing function of time given covariate. A kernel smoothed Grenander-type estimator for the non-increasing conditional density is used to construct the test statistics. We study the asymptotic normality of the test statistics and a bootstrap procedure is used to approximate the critical value of the test. The performance of the test is investigated through a simulation study, and we illustrate the practical use of the proposed method on a breast cancer dataset.

Keywords:
Isotonic estimation; Kernel smoothing; Mixture cure models; Sufficient follow-up; Survival analysis.

Robust estimation of occupation probabilities of latent multi-state processes

N. Maltzahn[1], G. D. Lukats[1], K. Røysland[1]

[1] Oslo Centre for Biostatistics and Epidemiology

E-mail: niklas.maltzahn@medisin.uio.no

Abstract:
Over the past three decades, scholars working in the fields of causal inference and missing data, have developed a variety of techniques to produce estimators with highly desirable robustness and efficiency properties. We apply these techniques, to derive various AIPW (augmented inverse probability weighted) estimators of occupation probabilities of a latent multi-state process under two levels of coarsening; right censoring and baseline exposure, allowing for both time constant and time dependent confounders. The key exchangeability assumption for identification is coarsening at random (CAR). The AIPW estimators are motivated from a different and arguably simpler identification result, than the common product integral representation of occupation probabilities. We investigate the performance of the estimators of occupation probabilities in a simulation experiment under different scenarios.

Keywords:
Occupation probabilities; Coarsening at random; Robust estimation.

Interventional dynamic updating of prognostic survival models in a pandemic environment

I. Prosepe[1], A. Kaal[2], C. Stark[1], J. Smit[3], E. Leegwater[4], C. van Nieuwkoop[2], S. Arbous[5], N. van Geloven[1]

[1] Department of Biomedical Data Sciences, Leiden University Medical Center, the Netherlands
[2] Department of Internal Medicine, Haga Teaching Hospital, the Netherlands
[3] Department of Intensive Care, Erasmus Medical Center, the Netherlands
[4] Department of Hospital Pharmacy, Haga Teaching Hospital, the Netherlands
[5] Department of Intensive Care, Leiden University Medical Center, the Netherlands

E-mail: i.prosepe@lumc.nl

Abstract:
Predictive algorithms are widely used in hospitals to aid patient management, but when treatment policies shift, they can quickly lose accuracy. While predictive models can be dynamically updated, these updates take time due to the need to gather sufficient new data. For instance, a clinical model's predictive accuracy may decline when a new treatment is introduced, requiring extensive sample sizes to re-estimate the effect of the predictors.

In this work, we delineate an alternative approach where external information on the effectiveness of the new treatment is used to estimate predictions under the upcoming intervention strategy. Predictions under interventions, which address "what if" scenarios by estimating outcomes under hypothetical interventions, have been used primarily for personalized decision-making. Here, we combine them with model updating strategies to create "interventional updates," aiming to reduce lag in dynamic model updates when new treatments are introduced in clinical practice.

For each treatment that changes in standard care—whether newly introduced, discontinued, or modified in use—there are two key time points: when the changes in care begin (e.g., when a new treatment starts being used) and when these changes stabilize (new treatment has become the new normal). This defines three distinct phases: the old standard care, a transitional period, and the new standard care. In the old and new standard care phases, standard updating techniques can be used. In the transitional period we propose to use interventional updating which incorporate evidence of treatment effect from external sources. We carefully write down how to extend known methods for dynamic updates for time-to-event prediction models, extending the most commonly known updating strategies: refitting, intercept recalibration and Bayesian updating methods.

We illustrate our methods using electronic health records from 4,064 patients hospitalized with COVID-19 in four Dutch hospitals in 2020 and 2021. A prediction algorithm was trained on first-wave data (February–July 2020) to estimate the 28-day risk of mortality or ICU admission from the time of hospital admission. We then compared the performance of "standard" and "interventional" model updates on second-wave data (August 2020–May 2021), assessing discrimination (c-index and AUCt), calibration (calibration intercept, slope and OE ratio), and overall performance (Brier score). For the "interventional" updates, we focused on the change in dexamethasone use, as this was identified as the most significant treatment shift in 2020. We incorporated evidence on dexamethasone's effectiveness from the RECOVERY trial.

Keywords:
Prediction models; Causal inference; Model updating; Covid-19.

Surviving your PhD: An analysis of time to completion data

A. M. Richardson[1]

[1] Australian National University, Australia

E-mail: alice.richardson@anu.edu.au

Abstract:
At Australian universities, PhD candidates are expected to submit their thesis within three to four years of starting their project. Extensions are possible, and were widely granted for students whose projects were affected by COVID. However, funding is tight, and University administrators are keen to know how long candidates are taking to submit their theses, and whether targeted support mechanisms such as statistics advice, intensive writing retreats, and the like make a measurable impact on time to completion.

This presentation will focus on a comprehensive analysis undertaken at the Australian National University to address these pressing questions. We will begin by outlining the structure of the PhD program and the nature of the data available to answer the research questions. This will include a detailed description of the various support mechanisms in place and the metrics used to measure their effectiveness.

While the survival models used to analyse the data will be off-the-shelf, the presentation will highlight the complexities encountered in understanding and interpreting the data. Two proposed solutions will be discussed in depth, revealing the challenges that analysts face when dealing with secondary analysis of administrative data in an educational context. These solutions will be evaluated in terms of their generalisability and scalability, as well as potential limitations.

The presentation will conclude by drawing connections between the research question, proposed solutions, and alternative analyses from within a causal framework. This will involve a critical examination of the assumptions underlying our approach and a discussion of how causal inference methods could potentially enhance our understanding of the factors influencing PhD completion times.

By sharing these insights, I aim to contribute not only to the technical discussion of survival analysis methodology but also to the broader conversation on improving PhD completion rates and the effective allocation of resources in higher education.

Keywords:
Causal models; Competing events; Higher education; Secondary data analysis.

References:

A. Kennard, A. Richardson, S. Rainsford, K. Hamilton, N. Glasgow, K. Pumpa, A. Douglas and G. Talaulikar (2024): Longitudinal frailty assessment in the prediction of survival among patients with advanced chronic kidney disease: a prospective observational cohort study. BMJ Open 14: e087189, https://doi.org/10.1136/bmjopen-2024-087189.

S. Das, B. Baffour and A. Richardson (2024): Trend estimation of sub-national level daily smoking prevalence by age and sex in Australia. Tobacco Induced Diseases 33, 45. https://doi.org/10.18332/tid/183804.

B. G. Heubeck, A. Richardson and G. Lauth (2024): Parent stress and social support in a randomized controlled trial of individual versus group parent training for children with HKD/ADHD. British Journal of Clinical Psychology, https://doi.org/10.1111/bjc.12483.

A ”what if”-interpretation of the Kaplan-Meier estimator and, in general, no such interpretation for competing risks

S. Schmeller[1], D. Fuerst[2], J. Beyersmann[1]

[1] Institute of Statistics, Ulm University, Ulm, Germany
[2] Institute of Clinical Transfusion Medicine and Immunogenetics Ulm, German Red Cross Blood Transfusion Service, Baden Wuerttemberg – Hessen, Ulm, and University Hospital Ulm, Germany and Institute of Transfusion Medicine, University of Ulm, Germany

E-mail: sandra.schmeller@uni-ulm.de

Abstract:
It should be well known that in a competing risk setting the use of the Kaplan Meier estimator counting only one type of event is biased for estimation of the cumulative incidence probability (Schuhmacher et al., 2016; Schmeller et al., 2023). Several analytical proofs of this exist which confirm examples from real data. Therefore, it is clear that the false-Kaplan-Meier is not an estimator for the cumulative incidence but it is questioned if this estimator still has a meaningful interpretation. The aim of this talk is to show a proof that the Kaplan-Meier estimator has a causal interpretation for the survival probability in a time to combined endpoint analysis that would have been observed if the random censoring had been avoided. However, the talk will show that this cannot be adapted to a competing risk setting and the Kaplan-Meier estimator counting only one type of event and censoring the other cannot be interpreted as the estimator for the cumulative incidence of the event of interest if the competing events had been avoided (the intervention distribution with intervention ”no competing event”). The reason is that the occurrence of the competing events are not independent. The simple proof is based on a simple causal graph and a straightforward application of the g-computation rule. It is conceptually related to but technically simpler than the recent argument by Young et al. (2020) however, without requiring that one cause of death precedes the other as in Young et al. (2020).

Keywords:
Kaplan-Meier estimator; Causality; Competing risks.

References:

M. Schumacher, K. Ohneberg and J. Beyersmann (2016): Competing risk bias was common in a prominent medical journal. J Clin Epidemiol, 80:135–136.

S. Schmeller, D. Fürst and J. Beyersmann (2023): Konkurrierende Risiken Modelle. In J. Gertheis, M. Schmid, and M. Spindler: Moderne Verfahren der Angewandten Statistik. Springer Spektrum, Berlin, Heidelberg, 2023.

J. G. Young, M. J. Stensrud, E. J. Tchetgen Tchetgen and M. A. Hernán (2020): A causal framework for classical statistical estimands in failure-time settings with competing events. Stat Med, 39(8):1199–1236.

A non-parametric proportional risk model to assess a treatment effect in an application to randomized controlled trials

L. Ameis[1], O. Kuss[2], A. Hoyer[3], K. Möllenhoff[1]

[1] Institute of Medical Statistics and Computational Biology (IMSB), Faculty of Medicine, University of Cologne, Cologne, Germany

[2] German Diabetes Center, Leibniz Institute for Diabetes Research at Heinrich Heine University Düsseldorf, Institute for Biometrics and Epidemiology, Düsseldorf, Germany

[3] Biostatistics and Medical Biometry, Medical School OWL, Bielefeld University, Bielefeld, Germany

E-mail: lucia.ameis@uni-koeln.de

Abstract:

Time-to-event analysis often relies on prior parametric assumptions or, if a non-parametric approach was chosen, Cox's proportional hazards model that is inherently tied to an assumption of proportional hazards. This limits the quality of the results in case of any violation of these assumptions. Especially the assumption of proportional hazards was recently criticized for being rarely verified. In addition, most interpretations focus on the hazard ratio, that is often misinterpreted as the relative risk and comes with the restriction of being a conditional measure. Our approach introduces an alternative to the proportional hazard assumption and allows for a direct estimation of the relative risk as well as the absolute measure of the number needed to harm, therefore provides the possibility of an easy and holistic interpretation.

In this talk, we propose a new non-parametric estimator to assess the relative risk of two groups to experience an event under the assumption that the risk is constant over time, namely the proportional risk assumption. Precisely, we first estimate the respective cumulative distribution functions of both groups by means of the Kaplan-Meier estimator and second combine their ratio at different time points to estimate the mean relative risk. We then combine the result with one of the estimated cumulative distribution functions to assess the number needed to harm. This offers the possibility to interpret the treatment effect solely based on a Kaplan-Meier estimator and offers a flexible alternative to Cox's model if the proportional hazard assumption is violated.

We demonstrate the validity of the approach by means of a simulation study and present an application to data from a large randomized controlled trial investigating the effect of dapagliflozin on all-cause mortality.

Keywords:

Time-to-event analysis; Risk; Number needed to treat; Treatment effect; Hazard ratio.

References:

L. Ameis, O. Kuss, A. Hoyer and K. Möllenhoff (2024): A nonparametric proportional risk model to assess a treatment effect in time-to-event data. Biometrical Journal, 66(4), 2300147.

Exhausting the type I error level in a group-sequential design with a closed testing procedure for progression-free and overall survival

M. F. Danzer[1], K. Rufibach[2], J. Beyersmann[3], R. Schmidt[1]

[1] University of Münster, Germany
[2] Merck KGaA, Darmstadt, Germany
[3] University of Ulm, Germany

E-mail: moritzfabian.danzer@ukmuenster.de

Abstract:
In the planning and analysis of clinical trials with time-to-event endpoints, multiple outcomes that reflect the course of disease can be of high importance. This applies in particular to oncological trials in which both overall survival (OS) and progression-free survival (PFS) may be used as confirmatory endpoints where appropriate. In such settings, which the health authorities refer to as multiple primary endpoints, control of the family-wise error rate (FWER) can be of great importance (FDA, 2017; EMA, 2017).

Recently, a planning approach has been proposed that exploits the relationship between PFS and OS to potentially gain efficiency in trial designs (Erdmann et al., 2023). In particular, it takes into account that the inflow of information for the endpoints may be of different speeds and that assuming proportional hazards for both endpoints simultaneously is not realistic. In such a design, the FWER can be controlled in a conservative way by splitting the overall significance level between the two endpoints in a weighted Bonferroni procedure.

In this work, we want to address the extent to which we can achieve uniform improvements to this approach. To do this, we identify two different methods that we can also combine with each other. First, we want to exploit the dependence between the endpoints in the calculation of rejection bounds for endpoint-specific log-rank tests. To do this, a model-free characterization of the joint distribution of the test statistics across different analysis times and endpoints will be crucial (Lin, 1991). Second, we want to show to what extent this dependence can also be exploited in the context of a closed testing procedure within a group-sequential design (Anderson et al., 2022).

While we will initially limit ourselves to calculations and simulations within our specific example of PFS and OS, the basic approach should also be applicable to other cases.

Keywords:
Family-wise error rate; Illness-death model; Multiple primary endpoints.

References:

FDA (U.S. Food and Drug Administration) (2017): Multiple endpoints in clinical trials: guidance for industry.

EMA (European Medicines Agency) (2017): Guideline on multiplicity issues in clinical trials.

A.Erdmann, J. Beyersmann and K. Rufibach (2023): Oncology clinical trial design planning based on a multistate model that jointly models progression-free and overall survival endpoints. arXiv preprint arXiv:2301.10059.

D. Y. Lin (1991): Nonparametric sequential testing in clinical trials with incomplete multivariate observations. Biometrika, 78(1):123-131.

K. M. Anderson, Z. Guo, J. Zhao and L. Z. Sun (2022): A unified framework for weighted parametric group sequential design. Biometrical Journal, 64(7):1219-1239.

Sample size calculation based on differences of quantiles from right-censored data

B. Farah[1,2,5], A. Latouche[2,4], O. Bouaziz[3], X. Paoletti[2,5]

[1] Université Paris Cité, France
[2] Insitut Curie, France
[3] Université de Lille, France
[4] Conservatoire National des Arts et Métiers, France
[5] Université de Versailles Saint-Quentin-en-Yvelines, France

E-mail: biafarah28@gmail.com

Abstract:
When evaluating treatment effect, it is common to rely on the hazard ratio, typically by using the ubiquitous Cox model. In the presence of right-censoring, the hazard rate can be easily estimated from the observed data which makes this model very appealing. In Randomized Clinical Trials (RCT), standard methods already exist to determine the sample size when the estimand is a hazard ratio (typically the hazard ratio for comparing the effect of two treatments) based on either the log-rank test or the Cox model. However, in cancer studies, some treatments may have a late effect and the proportional hazard assumption imposed by the Cox model is no longer verified. Thus, we would like to shift from the hazard ratio to the difference in quantile of failure time as estimand because:

1. It allows for quantifying different treatment effects across quantiles;
2. Quantile regression doesn't assume proportional risks, making it appropriate for analyzing delayed treatment effects associated with immunotherapy;
3. It offers a clinically interpretable way to measure the benefit of one treatment over another as a function of time.

Our goal is to propose a sample size formula for evaluating treatment effects by comparing pre-specified quantiles in each treatment group. A versatile method for testing equality of quantiles was proposed by Kosorok (1999), which allows to either test simultaneously different quantiles or to test the same quantile at different analysis times in a group sequential clinical design. This method requires an estimator of the density of the distributions at the quantiles, for which we propose a gaussian resampling method inspired by Lin et al. (2015). We studied the effect of the variance of the generated normal variables on the estimation of the density and showed that this parameter has an influence on the quality of estimation. As a result, we developed a method to choose this variance parameter from the data in an efficient way.

We propose an explicit expression for the power of the test which allows us to derive a formula for computing minimal sample size. Extensive simulation studies that compare the power of the proposed method with other approaches from literature are also presented.

Keywords:
Sample size; Censoring; Immunotherapy.

References:

M. R. Kosorok (1999): Two-sample quantile tests under general conditions. Biometrika 86(4), 909-921.

C. Lin, L. Zhang and Y. Zhou (2015): Conditional quantile residual lifetime models for right censored data. Lifetime data analysis 21, 75-96.

One-sample survival tests for non-proportional hazards in oncology clinical trials: A simulation study

C. Szurewsky[1], G. Yin[2], G. Le Teuff1[1]

[1] CESP, INSERM U1018, Université Paris-Saclay, France
[2] Department of Statistics and Actuarial Science, University of Hong Kong Pokfulam Road, Hong Kong

E-mail: chloe.szurewsky@gustaveroussy.fr

Abstract:
In recent years, many one and two-stage designs for single-arm trials with time-to-event outcome have been proposed as an alternative to the randomized clinical trial which may be unfeasible to design for rare diseases in oncology (pediatric studies or personalized medicine). These designs rely on the one-sample log-rank test1 (OSLRT) and its modified version2 (mOSLRT) for comparing the survival curve of an experimental group to that of an external reference group. These tests are developed under the proportional hazards (PH) assumption which may be violated particularly when evaluating immunotherapies. We propose to adapt the OSLRT and evaluate alternatives for situations where PH does not hold. We extended Finkelstein's score test1 developed under PH by using a piecewise exponential3 model with change-points (CPs) for early, middle or delayed effect. For crossing hazards, we use an accelerated hazards model. We extended the restricted-mean survival time (RMST) based test4 to single-arm trials and constructed a maximum combination (maxCombo) test5 combining the mOSLRT, early and delay score tests. The performances (type I error and power) of the developed tests are evaluated through a simulation study of a phase II single-arm trials with an accrual and a follow-up of 3 and 4 years, respectively. The survival times are generated using an exponential distribution assuming no sampling variability for the reference group and a piecewise exponential for the experimental group. We varied the sample size of the experimental group (from 20 to 200 patients), the exponential censoring rate (from 0 to 35%) and the relative treatment effect (hazard ratio from 0.5 to 1). For illustration, we applied these different approaches to pediatric trials for neuroblastoma. Simulations show that the score tests are more conservative than the mOSLRT. When the data generation matches with the model, the associated score test is the most powerful even when the CPs are misspecified. The RMST-based test is more powerful than mOSLRT just for an early effect with censoring rate less than 15%. The maxCombo test is conservative and has a higher power than mOSLRT with enough large sample size (n>50 or n>100) but less than the right score test under non PH. To conclude, the score tests are efficient when the approximate values of CPs are known and the maxCombo test is an alternative when the CPs are unknown. Further researches may be conducted to study the impact of the reference survival curve variability and its survival distribution.

Keywords:
Single-arm trials; One-sample log-rank test; Non-proportional hazards; Piecewise exponential model; Combination tests.

References:

D. M. Finkelstein, A. Muzikansky and D. Schoenfeld (2003): Comparing Survival of a Sample to That of a Standard Population. CancerSpectrum Knowledge Environment 95(19), 1434 39.

J. Wu (2014): A New One-Sample Log-Rank Test. Journal of Biometrics & Biostatistics 05(04).

C. Chu, S. Liu and A. Rong (2020): Study Design of Single-arm Phase II Immunotherapy Trials with Long-term Survivors and Random Delayed Treatment Effect. Pharmaceutical Statistics 19(4), 358-69.

P. Royston and M. Kb. Parmar (2013): Restricted Mean Survival Time: An Alternative to the Hazard Ratio for the Design and Analysis of Randomized Trials with a Time-to-Event Outcome. BMC Medical Research Methodology 13(1), 152.

T. G. Karrison (2016): Versatile Tests for Comparing Survival Curves Based on Weighted Log-Rank Statistics'. The Stata Journal: Promoting Communications on Statistics and Stata 16(3), 678–90.

A “double copula” model for semi-competing risks data

A. Dineva[1], O. Kuss[2], A. Hoyer[1]

[1] Biostatistics and Medical Biometry, Medical School OWL, Bielefeld University
[2] Institute for Biometrics and Epidemiology, German Diabetes Center, Düsseldorf

E-mail: antoniya.dineva@uni-bielefeld.de

Abstract:
Semi-competing risks describe the general setting in which the primary focus is modeling the age at a non-terminal event (e.g., disease occurrence) when the investigated subjects are also at risk for experiencing a terminal event (e.g., death). That is, the terminal event might censor the non-terminal event, but not vice versa. As a result, the observed ages at the two events within the same individual are correlated. The statistical approaches for this setting can be divided in at least two general groups: 1) copula-based models and 2) illness-death models.

A common approach in the first group involves modeling the dependency between the two events by a bivariate copula with two marginal distributions for age at disease onset and age at death. However, such approaches are limited in their ability to distinguish between the two different modes of mortality, the one with and the one without the disease.

We propose a “double copula” model, that estimates the three marginal distributions in the semi-competing risks framework: 1) age at disease onset, 2) age at death for individuals with the disease and 3) age at death for individuals without the disease. The model is defined based on two bivariate Clayton copulas modeling the joint distribution of age at disease onset with each of the two ages at death separately. We assume the marginals of the two mortalities to be Gompertz distributed, whereas for the disease onset a Weibull distribution is adopted. Model parameters are estimated by maximum likelihood, accounting for the complex censoring and truncation mechanisms in a cohort study. The likelihood function incorporates left truncation for both terminal and non-terminal events, reflecting delayed entry into the study. As in cohort studies the exact age at disease onset is only known to lie between two intermittent follow-up visits, we additionally consider interval censoring for the age of disease occurrence. We performed a simulation study, that demonstrated promising results in terms of accuracy and numerical robustness.

The model is illustrated using data from the Paquid study, a large cohort study on mental and physical aging with age at dementia onset describing the non-terminal event. Our model leads to plausible results, indicating that people with dementia diagnosis die earlier compared with people without dementia.

Keywords:
Competing risks; Copula model.

Incorporation of a mixture distribution on frailty regression model for clustered survival data

G. Kiprotich[1]

[1] Ludwig-Maximilians-Universität München, Germany

E-mail: Gilbert.Kiprotich@stat.uni-muenchen.de

Abstract:
In frailty models, the time-dependent correlation among observations is driven by the frailty distribution, which shapes how this dependence evolves. This paper introduces a new approach where the frailty is modeled using mixture distributions. The parameterization of the mixture distribution directly defines the mixing weights, and the model's closed-form Laplace transform allows for the calculation of Kendall's tau, a measure of dependence. We explore both parametric and semiparametric versions of this frailty model. Additionally, we present a hierarchical representation of the mixture, which simplifies the application of the expectation-maximization (EM) algorithm for parameter estimation. The effectiveness of the proposed model is demonstrated through Monte Carlo simulations and in an application to a cancer dataset. A comparative analysis with existing frailty models shows the strengths of our approach. The implementation of our methodology is available in the R package textttextrafrail.

Keywords:
Clustered survival data; EM algorithm; Finite mixture; Frailty.

Comparing a time-to-event endpoint in a two-arm trial investigating personalized treatment

M. Müller[1]

[1] DKFZ Heidelberg, Germany

E-mail: marilena.mueller@dkfz-heidelberg.de

Abstract:
We consider a two-arm randomized clinical trial in precision oncology with time-to-event endpoint. The control arm consists of standard of care (SOC) whereas patients in the treatment arm are offered personalized treatment. However, some patients in the treatment arm do not receive personalized targeted treatment as it is not available or patients do not consent to it. Instead they also receive SOC.

Intention-to-treat analysis involves comparing the outcomes of patients receiving a mixture of treatment and SOC to patients receiving solely SOC. It is proposed to divide the patients into groups based on whether they receive their intended treatment or not. In this mixture model, patients' progression is assumed to follow different conditional intensity in either group.

Specifically, we model the conditional intensity functions parametrically. A regression model is presented, estimation of the conditionally intensity for the Cox model with proportional hazards is enabled and various testing procedures are discussed. Maximum Likelihood estimation on the partial likelihoods of the components is used representing individual patients and both arms simultaneously. Counting process theory as well as martingale theory is used to develop suitable test statistics for various settings of interest. We conduct an in-depth simulation study to complement the theoretical results.

We propose guidelines on how to account for presence of mixtures as well as giving insight on when it is necessary or appropriate to apply a more rigorous model.

R-package discSurv: A toolbox for discrete time survival analysis

T. Welchowski[1,2], M. Berger[1], D. Köhler[1], M. Schmid[1]

[1] Institute for Medical Biometry, Informatics and Epidemiology, University of Bonn
[2] Psychological Methods, Evaluation and Statistics, Department of Psychology, University of Zurich

E-mail: t.welchowski@psychologie.uzh.ch

Abstract:
In discrete survival analysis one considers time-to-event responses, in which time is measured on an ordinal discrete time scale. This includes situations, where the outcome values are intrinsically discrete or a grouped version of underlying continuous event times. That framework allows easier interpretation of hazard rates, is able to handle ties and can be embedded with appropriate data preprocessing into a generalized additive model framework, which allows to take advantage of existing software (Tutz and Schmid, 2016). In this talk we give an overview of the R add-on package (Welchowski et al., 2024), which is designed as a toolbox to support all phases of applied discrete time analysis with conventient help functions. Specifically, the package supplies preparatory functions for appropriate re-shaping of data and fitting regression models with censored single-event, competing risks as well as subdistribution hazards (Berger et al., 2020). In addition, it includes methods for assessing calibration, discrimination (Schmid et al., 2018) and goodness of fit of discrete survival models.

Keywords:
Statistical software; Discrete time; Competing risks; Subdistribution; Calibration; Discrimination.

References:

M. Berger, M. Schmid, T. Welchowski, S. Schmitz-Valckenberg and J. Beyersmann (2020): Subdistribution hazard models for competing risks in discrete time. Biostatistics, 2(3):449-466.

M. Schmid, G. Tutz and T. Welchowski (2018): Discrimination measures for discrete time-to-event predictions. Econometrics and Statistics, 7(0):153-164.

G. Tutz and M. Schmid (2016): Modeling Discrete Time-to-Event Data. Springer Series in Statistics.

T. Welchowski, M. Berger, D. Koehler and M. Schmid (2024): discSurv: Discrete Time Survival Analysis. R package version 2.2.0, https://cran.r-project.org/web/packages/discSurv/index.html.

Extending the vertical model: An alternative approach to competing risks with clustered data

S. Battaglia[1], M. Fiocco[2], H. Putter[2]

[1] University of Palermo, Italy
[2] Leiden University, Netherlands

E-mail: salvatore.battaglia91@gmail.com

Abstract:
Vertical modelling (VM) offers an alternative approach to competing risks models, particularly when relaxing the proportional assumption or when facing missing causes of failure. It focuses on the joint probability, denoted as P(T, D), of the time to failure T and the cause of failure D, which is decomposed into (P(D|T)) and (P(T)). Both components rely on observable quantities: the total hazard and the relative cause-specific hazard, which can be easily estimated using multinomial regression and Cox proportional hazard models adding for covariates.

This paper introduces a novel approach by incorporating a random component in each part of the model to address unobserved heterogeneity in the presence of clusters. Data from the EMergenze-URgenze database, including Emergency Department (ED) access records from 63 Sicilian EDs in 2019, are used to analyze the risk of hospitalization or discharge during the length of stay once admitted to the hospital in a multi-center setting. The Cumulative Incidence Function (CIF) estimated from the vertical mixed model (VMM) is compared with that of the traditional competing risks frailty model.

The VMM with independent random effects represents the first attempt to extend VM and can be viewed as an alternative to the frailty competing risks model when computational time for estimation methods is prohibitive or when exploring different aspects of the data is desired.

Keywords:
Vertical model; Emergency department; Competing risks; Random effects.

References:

C. Celsa, G. Cabibbo, C. A. M. Fulgenzi, S. Battaglia, D. J. Pinato et al. (2024): Hepatic decompensation is the major driver of mortality in hepatocellular carcinoma patients treated with atezolizumab plus bevacizumab: The impact of successful antiviral treatment. Hepatology, 10-1097.

F. Vizzutti, C. Celsa, V. Calvaruso, M. Enea, S. Battaglia, F. Schepis et al. (2023): Mortality after transjugular intrahepatic portosystemic shunt in older adult patients with cirrhosis: a validated prediction model. Hepatology, 77(2), 476-488.

Patient disposition in clinical trials: Addressing competing risks with stacked probability and proportion plots

S. Doerken[1]

[1] DATAMAP GmbH, Germany

E-mail: sam.doerken@datamap.de

Abstract:
In clinical study reports, the very first analysis typically focuses on patient disposition, providing an overview of how many patients have completed, discontinued, or are ongoing in the trial. This analysis is critical for assessing if the trial can reach its goals and highlighting any potential issues. It is particularly important if the study is still ongoing. A comprehensive disposition analysis can thus be just as important as the patient outcome analysis.

Although disposition tables are typically displayed as simple frequency statistics, they can be enriched through time-to-event analyses or, even better, with time-to-event plots. However, the range of survival analysis plots is limited, and the standard Kaplan-Meier plot, while widely used, is often misapplied when competing risks are involved.

In my presentation, I illustrate that stacked probability plots are an effective alternative, providing a clear visual representation that addresses the issue of competing risks. I further argue that a straightforward stacked proportion plot that illustrates descriptive proportions over time is yet another, more pragmatic alternative; it is both very easy to interpret and also immune to competing risks, making it an ideal choice for conveying patient disposition in ongoing clinical trials.

Keywords:
Patient disposition; Time-to-event plots; Competing risks; Stacked probability plots; Stacked proportion plots.

Discrimination performance in illness-death models with interval-censored disease data

A. J. Rueten-Budde[1], M. Spreafico[1,2], H. Putter[2,1], M. Fiocco[1,2,3]

[1] Mathematical Institute, Leiden University, Leiden, The Netherlands
[2] Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, The Netherlands
[3] Trial and Data Center, Princess Máxima Center for Paediatric Oncology, Utrecht, The Netherlands

E-mail: m.spreafico@math.leidenuniv.nl

Abstract:
Illness-death models are used in survival analysis to study transitions between health states, particularly modelling progression from a disease-free state to illness and subsequently from illness to death. These models are frequently applied to clinical data to understand disease progression; however, in clinical practice, the exact timing of disease onset is often unobserved. Instead, this information may be interval-censored or unobserved due to death or censoring, which can bias estimates of disease incidence and regression coefficients. This challenge commonly arises in settings where diseases, like soft tissue sarcoma, can only be diagnosed at scheduled follow-up visits. For example, after initial surgery to remove a tumour, a patient may develop metastases that are only detected at these follow-ups – indicating the metastases developed sometime between the last negative and the first positive screening – resulting in interval-censored data for a binary time-dependent disease marker. Although ignoring interval-censoring introduces bias in estimating disease incidence and regression coefficients, its impact on model performance remains unclear.

This study investigates the effect of ignoring the interval-censored nature of observation times on model discrimination performance, measured in terms of time-specific Area Under the Receiver Operating Characteristic Curve (AUC) in both incident/dynamic and cumulative/dynamic definitions. Four different methods to estimate the illness-death model are compared: (1) the Cox model with disease state as time-dependent binary covariate (ignoring the interval-censored nature of the time-dependent covariate), (2) the piecewise-constant model accounting for interval-censoring estimated using the msm R-package, (3) the Weibull and (4) the M-spline models accounting for interval-censoring estimated using the SmoothHazard R-package. A simulation study is conducted considering several data scenarios simulated from illness-death models with Weibull transition hazards, considering different (i) sample sizes, (ii) types of death censoring, and (ii) follow-up visit intervals for observing the disease marker. Finally, the four methods are applied to a dataset of patients with high-grade soft-tissue sarcoma, and their performances are discussed.

The findings of this study suggest that, in the presence of interval-censored disease times, it is important to account for interval-censoring not only when estimating the parameters of the model but also when evaluating the discrimination performance of the disease.

Keywords:
Illness-death model; Interval-censoring; AUC.

Transitions, sojourns, and bias: Simulation insights for transplant strategies in leukemia

Y. Xu[1], U. Mansmann[1]

[1] Institut für Medizinische Informationsverarbeitung Biometrie und Epidemiologie, LMU Munich, Germany

E-mail: yujunxu@ibe.med.uni-muenchen.de

Abstract:
In pharmacoepidemiologic and surgical research, immortal time bias often arises in observational settings, where the waiting period for treatment is misclassified as treatment benefits (Lévesque et al., 2010). Cox-based methods that account for time-dependent factors, e.g. time-dependent Cox regression and landmark models, are widely used to address this issue (Gleiss et al., 2018). However, interventions such as surgery may substantially alter patients' risk profiles, eroding the common memoryless assumption on a single time scale for underlying hazards (Andersen et al., 2021). Further complicating this multistate scenario is the interplay of transitions across prognostic states and their sojourn times. For example, disease progression while waiting for a transplant may render a patient ineligible for treatment—an aspect often inadequately addressed in these models.

In this work, we conducted a phase III simulation study to enable neutral comparison (Heinze et al., 2024; Morris et al., 2019) through: (1) defining the estimand for assessing mortality under different transplantation strategies in acute myeloid leukemia patients; (2) generating data from time-continuous Markov and semi-Markov processes that incorporate realistic Weibull and log-logistic distributions with multiple time scales and intrinsic rules for treatment administration; (3) assessing bias and other performance measures for estimates derived from different analytical approaches, including time-dependent Cox regression, landmark models, multistate models, and the g-formula; and (4) reanalyzing aggregated acute myeloid leukemia data from the AMLSG and AMLCG cohorts based on our simulation results.

Keywords:
Stochastic process; Causal inference; G-formula; Multistate model; Immortal time bias.

References:

A. Gleiss, R. Oberbauer and G. Heinze (2018): An unjustified benefit: immortal time bias in the analysis of time-dependent events. Transpl Int, 31(2):125-130.

G. Heinze et al. (2024): Phases of methodological research in biostatistics-Building the evidence base for new methods. Biom J, 66(1):e2200222.

P. Kragh Andersen et al. (2021): Analysis of time-to-event for observational studies: Guidance to the use of intensity models. Stat Med, 40(1):185-211.

L. E. Lévesque et al. (2010): Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. Bmj, 340:b5087.

T. P. Morris, I. R. White and M. J. Crowther (2019): Using simulation studies to evaluate statistical methods. Stat Med, 38(11):2074-2102.

Poster list

Methods for analyzing multiple time-to-event endpoints in randomized clinical trials: A comprehensive overview

D. Alidan[1], A.-K. Ozga[1]

[1] University Medical Center Hamburg-Eppendorf, Germany

E-mail: d.alidan@uke.de

Abstract:
In clinical trials time-to-event analysis is crucial for assessing treatment efficacy and patient outcomes, e.g. time to myocardial infarction, time to hospitalization, or time to death in cardiovascular trials. Often thereby only the time to the first occurring event is considered and time-to-first-event methods, such as the Cox proportional hazards model and Kaplan-Meier estimate, are applied. However, they inherently overlook subsequent or competing events and fail to fully capture the complex nature of disease progression, potentially leading to incomplete or biased insights into treatment efficacy.

This research highlights a critical gap in conventional methods: their inability to differentiate between the clinical relevance of non-fatal and fatal events, and their neglect of the correlation and competition between multiple event types. For instance, myocardial infarction may heighten the risk of future events, including subsequent infarctions or death, creating a complex interrelationship that traditional models fail to address. This not only limits the insights drawn from such trials but also undervalues the comprehensive patient experience.

To overcome these limitations, we explore advanced methods for recurrent and competing events, including Cox based models for recurrent event analysis, multistate models, or flexible parametric models. Through a thorough comparison of semi-parametric, non-parametric, and parametric methods, we aim to provide applied researchers with a clear understanding of the advantages, assumptions, and limitations of each approach. Importantly, we examine which methods best align with the complex nature of real-world clinical trial data, particularly in randomized controlled trials comparing two treatment groups.

The novelty of our work lies not only in providing a systematic overview but also in applying these methods to a real-world cardiovascular dataset, showcasing how multiple event types (e.g., hospitalization, stroke, and death) can be more accurately analyzed.

By addressing these challenges, this study paves the way for more sophisticated analyses that can better inform treatment decisions and improve patient care. We aim to demonstrate how embracing these advanced methods can unlock deeper, clinically relevant insights that were previously inaccessible using traditional time-to-first-event techniques.

Reconstructing survival curves: Using imputation strategies to construct Kaplan-Meier estimates with no or limited data on survivors

L. Berchtold[1], T. Gorlia [2], M. Geurts [3], M. Weller[4], M. Preusser[1], F. König[1]

[1] Medical University of Vienna, Austria
[2] EORTC Headquarter, Brussels, Belgium
[3] Erasmus MC Cancer Institute, Rotterdam, the Netherlands
[4] University Hospital and University of Zurich, Zurich, Switzerland

E-mail: luzia.berchtold@meduniwien.ac.at

Abstract:
Secondary use of clinical trial data has become increasingly valuable in medical research, providing new insights without the expense of conducting additional trials. However, legal, ethical, and proprietary restrictions often limit access to datasets, e.g., by providing only access to the data of deceased patients. Analysing restricted datasets only carries a risk of bias since the results depend on the proportion of subjects missing. This creates a major challenge when interpreting such survival analyses. Conventional methods like multiple imputation cannot be applied as they would require information on missing patients, making them unsuitable in cases where no data from surviving patients is available.

Motivated by a real case study, we conducted a simulation study to explore the strength of the bias in different settings. The main methods of analysis include Kaplan-Meier survival curves, log-rank tests, and Cox regression models to estimate and compare survival probabilities. We explored different imputation strategies for censored patients when using other publicly available information, e.g., minimum and maximum censoring times. In the context of a randomized clinical trial we will discuss which additional information and assumptions will be needed to perform different imputation strategies. The impact of various strategies on operating characteristics such as type 1 error rate, power, and bias were evaluated. All analyses were conducted in R using the ”survival” package.

We show the severe limitations of performing an analysis on restricted datasets only. Kaplan-Meier estimates for the groups will be closer together, resulting in a loss of power. For instance, even in a scenario with only 10% administrative censoring and a true underlying hazard ratio of 0.7 and a sample size of 300, the power will be decreased from 84% when using the complete datasets to 39% power in the restricted setting. As censoring increases, differences in the Kaplan-Meier plot become harder to detect, and both the survival function and estimates like median overall survival (OS) are underestimated.

In conclusion, restricting survival analysis to deceased patients introduces a significant bias, including underestimated survival functions and reduced statistical power, especially with higher proportions of censoring. When some information about the missing patients and censoring process is available, even a uniform imputation strategy greatly improves estimation accuracy. However, this information is often not available when analyzing non-randomized variables and the imputation strategy is not applicable. In these cases a worst case analysis should be carried out.

Comparison of the prognostic performance of machine learning algorithms on gene expression data in acute myeloid leukemia

A. Blanda[1], S. Pizzamiglio[1], S. Sangaletti[2], P. Verderio[1]

[1] Unit of Bioinformatics and Biostatistics, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy
[2] Immunology Unit, Department of Experimental Oncology, Fondazione IRCCS Istituto Nazionale Tumori, Milan, Italy

E-mail: adriana.blanda@istitutotumori.mi.it

Abstract:

Bckground: Acute Myeloid Leukaemia (AML) is a heterogeneous hematologic malignancy characterized by the clonal expansion of abnormal myeloid progenitor cells. Despite advances in treatment, the prognosis for patients remains variable with an estimated 5-year overall survival (OS) of approximately 30%. In recent years, gene expression analysis has emerged as a powerful tool for identifying novel biological markers that can aid in predicting patient outcomes. The main objective is to assess and compare the prognostic performance of different Machine Learning (ML) algorithms.

Methods: This study utilized a dataset comprising gene expression profiles from approximately 20,000 genes in 457 AML patients aged under 60, generated through RNA sequencing. The used dataset is available on the Gene Expression Omnibus platform (GSE6891). The outcome variable is OS defined a time from diagnosis to death. The median follow-up time is 44 months range [18-68] with a total of 290 events. The dataset was splitted in training (70%) and testing (30%) sets. To begin filtering, in the training set, a univariate Cox proportional hazards model was performed, followed by p-value adjustment using the False Discovery Rate (FDR) method (P<0.05). As feature selection methods, Lasso, Elastic-net, Forward and Backward approaches were applied. Subsequently, supervised algorithms such as Survival Random Forest, mBoost, Support Vector Machine and, Bayes Cox Models were implemented to improve the performance of models thanks to cross-validation and parameter tuning. The Harrel's Concordance index and its confidence interval was used to compare the prognostic performance.

Results: The candidate genes based on FDR-adjusted p-values were 83. After feature selection, the best model in terms of Harrell's C-index was a model with 7 genes, obtained by backward selection procedure C:0.68 [95% CI: 0.56–0.78]. The algorithm with the best performance was mBoost, achieving a C-index of 0.68 [95% CI: 0.55–0.79], followed by the Survival Random Forest model with a C-index of 0.65 [95% CI: 0.51–0.76]. In contrast, both the SVM and Bayes Cox models had lower C-indices of 0.56, with CIs of 0.42–0.68 and 0.48–0.52, respectively.

Conclusions: These results showed that the mBoost algorithm was the best for this dataset. To obtain more robust and generalizable results, we plan to evaluate the prognostic performance of the signature on an external dataset. Furthermore, the present analysis illustrates a process that integrates standard statistical method for selecting features, combined with ML algorithms to build the final model.

Keywords:
Harrell's C; High dimensionality; Performance; Survival ML algorithm; Pipeline.

Building risk prediction models by synthesizing national registry and prevention trial data

O. Chernova[1], D. Ankerst[1]

[1] Technical University of Munich, Germany

E-mail: oksana.chernova@tum.de

Abstract:
Models that predict cancer incidence aid developing personalized screening regimens and models that predict outcomes after a cancer diagnosis are used to individually tailor treatments. Developing risk prediction models requires specifying a set of predictors at the baseline of the prediction and a projection period of interest, commonly five-years, thus requiring cohorts with adequate follow-up. In recent decades, there has been a notable increase in the accessibility of extremely large datasets for research purposes. These expansive repositories encompass diverse data sources such as population-based census records, disease registries, healthcare databases, and collaborative initiatives from numerous individual studies. The notable advantage of these extensive datasets is their impressive sample size, which enhances their utility for research endeavors. The first breast incidence cancer risk prediction model, the Breast Cancer Risk assessment Tool (BCRAT, also known as The Gail Model) was built by combining odds ratio estimates for risk factors approximating relative risks from a case-control study with age-specific breast cancer incidence and competingmortality data on US women from the Surveillance, Epidemiology and End Results (SEER) registries. The statistical method of combining has since been extended to develop similar clinical risk models for projecting cancer risk in special populations as well as for different cancers; see for example, Gail et al. (1989, 2007), Chen et al. (2006), Freedman et al. (2009) and Pfeiffer et al. (2013). The method has not been applied to build risk models for the projection of five-year prostate cancer risk to the knowledge of authors of this study. Gelfond et al. (2022) developed a five-year prostate cancer risk model using two widely known prostate cancer prospective screening studies that completed in the early 2000s to be used in this article. Their model did not incorporate contemporary incidence rates from SEER. The purpose of this study was to provide the statistical methodology and corresponding R implementation for building pure risk prediction models merging cohort and SEER incidence data, exemplified through the development of a five-year prostate cancer risk prediction tool.

Keywords:
Information synthesis; Cancer registries; Prostate cancer risk prediction.

References:

J. A. Gelfond, B. Hernandez, M. Goros et al. (2022): Prediction of future risk of any and higher-grade prostate cancer based on the PLCO and SELECT trials. BMC Urol 22, 45.

R. M. Pfeiffer, Y. Park, A. R. Kreimer, J. V. Jr. Lacey, D. Pee, R. T. Greenlee, S. S. Buys, A. Hollenbeck, B. Rosner, M. H. Gail and P. Hartge (2013): Risk prediction for breast, endometrial, and ovarian cancer in white women aged 50 y or older: derivation and validation from population-based cohort studies. PLoS Med., 10(7):e1001492.

R. M. Pfeiffer and M. H. Gail (2018): Absolute Risk: Methods and Applications in Clinical Management and Public Health. CRC Press.

R. M. Pfeiffer, Y. Chen, M. H. Gail and D. P. Ankerst (2022): Accommodating population differences for validating risk prediction models. Statistics in Medicine, 41(24):4756-80.

Extending landmarking to mixture cure models with time-varying covariates

M. Cipriani[1,2]

[1] Sapienza University of Rome
[2] Leiden University, Italy

E-mail: marta.cipriani@uniroma1.it

Abstract:
Mixture cure models are commonly used in survival analysis when a subset of individuals is considered cured and no longer at risk for the event of interest. A key challenge in such models arises when covariates are time-varying, requiring dynamic updates in both the cure probability and the survival prediction for the uncured population.

Landmarking, a dynamic approach that updates survival predictions at specific time points, has been proposed for mixture cure models (Shi and Yin, 2017). However, existing landmarking methods, such as the last observation carried forward (LOCF) approach, often fail to fully capture the complexity of longitudinal data.

To address this limitation, we propose a novel framework that extends hierarchical (mixed-effects) models to incorporate time-varying covariates in both the incidence (cure probability) and the latency (survival for the uncured) components of a mixture cure model. Hierarchical models, also known as multilevel or mixed-effects models, allow for random effects accounting for the individual-specific unobserved heterogeneity over time.

In the proposed framework, we model longitudinal covariate trajectories using a mixed-effects model to estimate individual-level random effects. These random effects are then integrated into the mixture cure model, providing a robust way to predict survival outcomes based on dynamically updated covariate information.

Our approach integrates the most recent covariate information into both the incidence (cured) and the latency (uncured) components of the mixture cure model, enabling more accurate, dynamic predictions for individual patients. This approach addresses the uncertainty in covariate estimation and adjusts survival predictions in real time as new information becomes available.

We show the features of this approach through a simulation study to compare our extension to current landmarking methods. The simulations help to illustrate the advantages of using individual-specific random effects in improving prediction accuracy, reducing bias, and providing more flexible modelling of time-varying covariates. Additionally, we apply our method to a real-world dataset including liver-transplanted patients, highlighting the improved performance of dynamic survival predictions in complex longitudinal settings.

Investigating the most suitable modelling framework to predict long-term restricted mean survival time and life expectancy

H. L. Cooper[1], M. Rutherford[1], S. Booth[1]

[1] Biostatistics Research Group, Department of Population Health Sciences, University of Leicester

E-mail: hc381@leicester.ac.uk

Abstract:

Background: Survival statistics provided at fixed points after diagnosis are often misinterpreted, with difficulties stemming from understanding the scale of the measures, and the fact that various estimands are frequently used interchangeably despite having different interpretations. Life expectancy estimates have been shown to be more easily understood however are not commonly used in practice as they often require extrapolation. We sought to explore: the best modelling framework for providing this extrapolation, how much follow-up is required, and under which circumstances valid estimates of Life Expectancy/40-year Restricted Mean Survival Time (LE/40-year RMST) can be obtained.

Methods: Data from the Surveillance, Epidemiology and End Results program (SEER) involving nine cancer registries across the USA was used to develop two flexible parametric models of varying complexity (Model 1: Solely main effects, Model 2: Main effects, interactions and time dependent effects) for patients diagnosed with colon cancer. Complete case analysis was conducted resulting in 35,903 colon cancer patients under investigation. The two models were compared within an all-cause, cause-specific and relative survival framework and the amount of follow up time used to develop the models was varied (2-, 3-, 5-, 10-, 20-years). 40-years of follow up was also inspected to assess model fit with all available data. To assess the accuracy of the extrapolations, marginal predictions were made from each of the statistical models and for each follow-up scenario across 30 covariate groupings based on age, sex and stage of tumour. LE/40-year RMST and 40-year survival probability were calculated. Estimates were then compared to the observed values based on Kaplan-Meier estimates.

Results: Reasonable accuracy of LE/40-year RMST for all covariate groups was generally demonstrated using 10+ years follow up. The relative survival framework demonstrated to be generally most suited to predicting LE/40-year RMST. Increasing model complexity improved predictions for younger patients however limited difference was demonstrated for higher risk patients (older and/or distant cancer).

Conclusion: This study demonstrated increasing model complexity is not always paramount to estimate LE/40-year RMST with recommended model complexity dependent on the patient's risk. Results demonstrate LE/40-year RMST estimates do not require long follow up, where often 10 years of follow-up is sufficient. Findings offer the ability to reasonably estimate LE for colon cancer patients of varying risk. Further work to understand the generalisability of this study to other cancers is needed. Likewise, internal and external validation should be conducted to support findings.

Keywords:

Life expectancy; Flexible parametric survival models; Prognostic model; Cancer.

Introducing a flexible model for regression models with a left-censored response and covariate

I. De Batselier[1], R. Braekers[1]

[1] Hasselt University, Belgium

E-mail: inez.debatselier@uhasselt.be

Abstract:
In most studies with left-censored data, the focus lies on regression models where only the response is left-censored. In this work we also add a left-censored covariate, which complicates the situation. In this situation people often resort to naive methods, such as a complete case analysis or a limit of detection substitution. To step away from these methods, Thao et al. (2021) proposed a method in which a parametric assumption is made about the distribution of the covariate. With our new method we want to avoid this parametric assumption. Instead of implying this assumption on the covariate, we estimate its distribution via a flexible model, namely through introducing a piecewise exponential distribution with fixed cut-off points. This piecewise exponential distribution offers robustness against misspecification of the covariate's distribution. In our methodology we propose a one-stage and a two-stage estimation. In the one-stage estimation all the parameters are estimated simultaneously. While in the two-stage estimation the parameters from the piecewise exponential distribution are estimated first and then plugged into the likelihood to estimate the rest of the regression parameters. The one-stage estimation is theoretically more feasible, but may introduce dimensionality problems due to the many parameters to be estimated. This dimensionality problem can be reduced by the two-stage estimation. Via simulations we show that our new methodology produces unbiased estimates for medium-sized datasets and fairly large amounts of censoring in the covariate and the response. A comparison of our method with other methods, via simulations, shows that it outperforms the others.

Keywords:
Regression models; Left-censored covariate; Flexible parametric model; Piece-wise exponential.

References:
T. Thao, A. Steven, A. Marc, M. Kirsten and H. Niel (2021): Measuring association among censored antibody titer data. Statistics in Medicine, 40(16), 3740-3761.

An R function for data preparation for an acyclic multistate model with non-ordered intermediate states

L. Del Castello[1], D. Bernasconi[1,2], L. Antolini[1], M. G. Valsecchi[1,3]

[1]1Bicocca Bioinformatics Biostatistics and Bioimaging Centre - B4, School of Medicine and Surgery, University of Milano-Bicocca, Monza, Italy

[2] Functional Department for Higher Education, Research, and Development, ASST Grande Ospedale Metropolitano Niguarda, Milan, Italy

[3] Biostatistics and Clinical Epidemiology, Fondazione IRCCS San Gerardo dei Tintori, Monza, Italy

E-mail: l.delcastello@campus.unimib.it

Abstract:

Background: The package "mstate" is a popular R library to perform multistate analysis (Putter et al. 2007). The function "msprep" provides a simple tool to handle data preparation where a wide-format dataset (one row per patient) must be transformed in long-format (one row for each possible transition). However, "msprep" allows only for triangular transition matrices corresponding to irreversible acyclic Markov processes.Sometimes this is too restrictive. Hazard et al. (2020) used multistate models to analyze care pathways of patients admitted to ICU for severe COVID-19. Patients were allowed to move back and forth between non-absorbing states. Thus, the authors developed a dedicated R function for data preparation which overcomes the limits of "msprep".

Methods: A further complexity, not present in the application by Hazard, could be the case of an acyclic model with no defined order between intermediated states. In our application, patients affected by cardiogenic shock are admitted to ICU and then may be treated at different times with none, one, two or three mechanical circulatory supports, namely IABP, Impella or ECMO. A patient may be treated with any of these devices and there is no specific order between treatments. However, no patient can be treated twice with the same device (e.g., after IABP one could receive another device but then cannot go back to IABP). The final absorbing state is a composite endpoint (first event between LVAD implantation, heart transplant or death). In this context, transition to a certain intermediate state blocks the possibility of future back-transitions to that state.

Results: We developed an R function allowing for data preparation in this situation. The proposed function transforms a dataset from wide to long format that can be passed to the function "msfit". The new dataset will be prepared in order to cause the contribution to the risk sets in time to be consistent with the dynamic update of possible transitions. This is achieved through the definition of a new matrix, called "blockmat", in which the paths that cannot be observed are declared blocked. Without the specification of the blockmat, some subjects would be considered at risk of presenting also transitions which they can never travel. In practice, this would result in the presence of additional unnecessary rows in the long dataset that will cause an underestimation of transition hazards.

Conclusion: We created an R function allowing for data preparation in case of an acyclic multistate model with non-ordered intermediate states.

Keywords:

Multistate; Transition matrix; Data preparation; R function; mstate package.

References:

H. Putter, M. Fiocco and R.B. Geskus (2007): Tutorial in biostatistics: competing risks and multi-state models. Statistics in Medicine, 26(11), 2389-430, doi: 10.1002/sim.2712, PMID: 17031868.

D. Hazard, K. Kaier, M. von Cube, M. Grodd, L. Bugiera, J. Lambert and M. Wolkewitz (2020): Joint analysis of duration of ventilation, length of intensive care, and mortality of COVID-19 patients: a multistate approach. BMC Med Res Methodol., 20(1):206, doi: 10.1186/s12874-020-01082-z, PMID: 32781984, PMCID: PMC7507941.

Principled estimation and prediction with competing risks: A Bayesian nonparametric approach

C. Del Sole[1], A. Lijoi[2], I. Prünster[2]

[1] University of Milan - Bicocca, Milan (Italy)
[2] Bocconi University, Milan (Italy)

E-mail: claudio.delsole@unimib.it

Abstract:
Competing risks occur in survival analysis when multiple causes of death are present. They play a prominent role in several domains extending beyond biostatistics to encompass epidemiology, actuarial sciences and reliability theory. In this contribution, a multi-state modeling framework to competing risks is adopted; we introduce a flexible nonparametric prior, specified in terms of hierarchical completely random measures, for the transition rates, inducing dependence among the different sources of risk. The marginal distribution of the data and of a latent random partition, which admits a characterization in terms of a variant of the Chinese restaurant franchise process, is determined. Leveraging these distributional results, we are able to evaluate the predictive probability that a future event is of a specific type (e.g. death from a particular cause), as a function of the time at which the event occurs. The resulting functions, derived on sound principles, represent a major innovation in the literature and are termed prediction curves. In addition, we characterize the posterior distribution of the hierarchical random measures and provide estimates for the survival function, and for the cause–specific incidence and subdistribution functions, conditionally on such latent partition structure. Both a marginal and a conditional sampling algorithms for posterior inference are devised; the performance of the model and the effectiveness of the two algorithms are then assessed by means of a simulation study, in which the proposed model is also compared non-hierarchical counterpart, modelling transition rates independently for each source of risk. Finally, some applications to clinical datasets are presented.

Keywords:
Bayesian nonparametrics; Competing risks; Hierarchical processes; Predictive inference; Random measures.

Enhancing healthcare understanding from clinical routine data by simplifying the representation of treatment pathways

M. Farhadizadeh[1], Z. Lange[2], A. Sigle[3], H. Dette[2], H. Binder[4], N. Binder[1]

[1] 4 Institute of General Practice/Family Medicine, Faculty of Medicine and Medical Center, University of Freiburg, Freiburg, Germany
[2] Department of Mathematics, Ruhr University Bochum, Bochum, Germany
[3] Department of Urology, Faculty of Medicine, Medical Center, University of Freiburg, Freiburg, Germany Center, University of Freiburg, Freiburg, Germany
[4] Institute of Medical Biometry and Statistics, Faculty of Medicine and Medical Center, University of Freiburg, Freiburg, Germany

E-mail: maryam.farhadizadeh@uniklinik-freiburg.de

Abstract:

Background:

In clinical routine care data, patients' pathways often exhibit heterogeneity, encompassing the order in which a wide variety of interventions occur in addition to varying sequence length. This presents a challenge in understanding and managing healthcare paths effectively, e.g., when using multistate models, which can only deal with a limited number of states. Therefore, there's a critical need to identify and group similar pathways within this diverse landscape (Binder et al., 2022; Möllenhoff et al., 2024).

Methods:

To identify typical treatment pathways for inpatient stays, we have developed an algorithmic solution that uses coded clinical data, i.e., diagnoses and procedures. This approach helps to visually represent the data in treatment path diagrams. Specifically, the algorithm detects important procedures by calculating their importance based on patient counts and node significance, comparing them to predefined thresholds. Additionally, the algorithm groups less important procedures to put the focus on essential components of these pathways.

Results:

We demonstrate our algorithm using clinical routine data from prostate cancer patients receiving radical prostatectomy in the Department of Urology at the Medical Center, University of Freiburg. To further explore the efficacy of our approach in simplifying the representation of treatment pathways, we evaluate it through an extensive simulation study. This involves varying pathway similarities, i.e., different levels of heterogeneity of interventions, and sequence lengths. We find that a representation with a manageable number of typical pathways can be obtained, and characterize the sensitivity with respect to different tuning parameters of our algorithm.

Conclusion:

This method simplifies the identification and visualization of common healthcare trajectories, aiding in understanding patient paths and informing healthcare decisions. Improving our understanding of these pathways can elevate clinical care standards and enhance health outcomes, in particular by enabling subsequent analysis with multistate models and statistical tests.

Keywords:
Treatment pathways; Clinical routine data; Similarity.

References:

Binder, Möllenhoff, Sigle and Dette (2022): Similarity of competing risks models with constant intensities in an application to clinical healthcare pathways involving prostate cancer surgery. StatMed, 41(19), 3804-3819.

Möllenhoff, Binder and Dette (2024): Testing similarity of parametric competing risks models for identifying potentially similar pathways in healthcare. arXiv:2401.04490v1.

Does a SARS-CoV-2 infection increase the risk of dementia? An application of causal time-to-event analysis on real-world patient data

J. Guski[1], H. Fröhlich[1,2]

[1] Fraunhofer-Institut für Algorithmen und Wissenschaftliches Rechnen SCAI, Germany
[2] Universität Bonn, Germany

E-mail: jannis.guski@scai.fraunhofer.de

Abstract:
Despite the relative recency of the global pandemic, there is evidence that a SARS-CoV-2 infection may act as a catalyst in the gradual manifestation of neurodegenerative diseases. For example, comparisons of biomarkers before and after an infection revealed signs of changes in brain structure, neuroinflammation and disruptions of the blood-brain barrier.

As part of the EU-funded COMMUTE project, we deploy and adapt time-to-event models on real-world data to investigate whether a SARS-CoV-2 infection increases the risk or accelerates the process of developing Alzheimer's or Parkinson's disease, and which factors from a patient's history may play a role in this assumed co-pathology.

Any findings are potentially affected by baseline confounding, dependent censoring and death from any cause as a competing risk. These problems may be amplified by the problem that control patients have to be sampled from before the pandemic because presumably, almost everyone has had a documented or undocumented SARS-CoV-2 infection at some point after 2020.

In a first step, we estimate average exposure effects of COVID-19 on event risks with statistical models that address the potential sources of bias via Targeted Maximum Likelihood Estimation (TMLE). Predicted effects undergo a detailed sensitivity analysis with respect to different datasets, alternative control group design and dataset stratification (e.g., by infection wave, age and sex).

Furthermore, causal machine learning models are trained that will be able to make individualized risk and exposure effect predictions for new patients. These models may derive encodings of structured data from electronic health records, e.g., via transformer architectures.

At the SAfJR, we would like to present preliminary results from our medical application of causal inference in time-to-event analysis, and discuss with the expert audience how our methodology may be further refined.

Keywords:
Causal time-to-event analysis; Real-world data; Co-pathology; Machine learning.

Effective sample size for Cox models: A measure of individual uncertainty in survival predictions

T. Hackmann[1], D. Thomassen[1], S. le Cessie[1,2], H. Putter[1], E. W. Steyerberg[1,3], L. C. de Wreede[1]

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

E-mail: t.hackmann@lumc.nl

Abstract:

Background/Introduction: Clinical prediction models are becoming increasingly popular to support shared decision-making. Most prediction models prioritize the accurate estimation and clear communication of point predictions. Uncertainty around the point prediction may be expressed by confidence intervals or left out altogether. To present prediction uncertainty in an intuitive way, the concept of effective sample size has recently been introduced for linear and generalized linear models, but not yet for the Cox model. (Thomassen et al., 2024) Our goal is to provide estimates of the effective sample size for individual predictions based on a Cox model.

Methods: Effective sample size is defined as the hypothetical sample size of patients with the same characteristics (with respect to the model) as a new patient whom the prediction is for, such that the variance of the outcome in that sample would be the same as the prediction variance. It can be calculated as a ratio of the outcome variance conditional on the predictor values to the prediction variance. We estimate the effective sample size for Cox model predictions and investigate the behaviour of this estimate in an illustrative clinical data set of colon cancer patients and through simulations.

Results: The variance of a prediction based on a Cox model depends on the variance of the estimated coefficients and variance of the baseline hazard, which is impacted by censoring. This variance can readily be estimated asymptotically based on the delta method, or using resampling-based methods such the bootstrap. We will show the behaviour of effective sample size for patients as a function of follow-up time. Patients who have covariate values close to the population mean have a higher effective sample size, while patients with rare covariate combinations may have a very small effective sample size, and thus high uncertainty.

Conclusions: Effective sample size can express the uncertainty of predictions for individual patients from a Cox model. Future studies should clarify its role to communicate uncertainty around the point prediction of survival probabilities and a possible role in model building.

Keywords:
Prediction modelling; Uncertainty quantification; Risk communication; Survival analysis.

References:

D. Thomassen, S. le Cessie, H. C. van Houwelingen and E. W. Steyerberg (2024): Effective sample size: A measure of individual uncertainty in predictions. Statistics in Medicine, 43(7), 1384-1396, https://doi.org/10.1002/sim.10018, Funded by the European Union under Horizon Europe Work Programme 101057332.

Imputation free deep survival prediction using conditional variational autoencoders

N. Hong [1,2,3], C. Yau [1,2]

[1] University of Oxford
[2] Health Data Research UK
[3] The Alan Turing Institute

E-mail: natalia.hong@keble.ox.ac.uk

Abstract:
The availability of Electronic Health Records (EHR) offers an opportunity to develop risk prediction tools to support clinical decision making. Yet, EHRs are generated from active clinical processes, and unlike data from controlled studies, lack complete records as only information deemed necessary for patient management is captured. This selective capture implies that apparent missingness can itself be informative for risk prediction. Crucially, missing data occurs both during model training and deployment, which is challenging for algorithms that rely on fixed-size inputs. Effective predictive tools must accommodate this variability in inputs to ensure the optimal use of data available to clinicians at the point of care.

Missing data is commonly addressed through imputation, where missing entries are "filled-in" before being passed to the model. This approach requires extra validation of the imputation model and relies on untestable assumptions about the missingness mechanism, potentially leading to unreliable predictions. Imputation can also overlook the informative nature of missingness as models cannot distinguish between imputed and observed entries. Alternatively, the Missing Indicator Method introduces missingness indicators into the model, but risk overfitting to specific patterns of missingness. The sharing pattern submodel (SPS) approach builds a separate model for each missing pattern while encouraging information sharing across submodels, without making missingness mechanism assumptions. However, it suffers from combinatorial inefficiencies due to the exponential growth in the number of patterns (2^p) and is limited to linear models.

We introduce an imputation-free framework that employs Conditional Variational Autoencoders (VAEs) jointly trained with any deep survival model to predict risk using incomplete EHRs. Our approach leverages VAE's ability to learn the distribution of missing patterns within a latent space to capture similarities across patterns. The learned latent embedding is fed directly into the deep survival model, enabling non-linear modelling while avoiding the combinatorial inefficiencies of SPS. We demonstrate our proposed framework with the deep survival model DeSurv, through simulation studies and two retrospective cohorts from the Clinical Practice Research Datalink primary care database. Our results show that the proposed framework is more robust in generalising to unseen or rare missingness patterns, with improved performance according to calibration-based survival metrics. The incorporation of a variational structure allows the model to decouple the learning of data and missingness, offering a more nuanced understanding of how missingness influences predictions. This framework provides a practical, consistent approach to handling missing data across development, validation, and deployment stages, all while maintaining strong performance.

Keywords:
Missing data; Deep survival analysis.

Planning early-phase clinical trials in oncology: A comprehensive simulation approach for response, progression-free survival, and overall survival

U. Ippagunta[1,2], H. Götte[1]

[1] Merck Healthcare KGaA, Germany
[2] Universität Regensburg, Germany

E-mail: udeernaippagunta@gmail.com

Abstract:
In oncology, clinical trials rely on endpoints such as Objective Response, Progression-Free Survival (PFS), and Overall Survival (OS) to assess therapeutic efficacy (Delgado and Guddati, 2021). These endpoints are interdependent; for instance, a patient must be alive and progression-free to qualify as a responder. Their evaluation varies across follow-up times, with response typically assessed before PFS and OS within a trial. Ignoring these dependencies may affect decision-making accuracy.

To address this, we aim to model these endpoints simultaneously to support early-phase clinical trials and reduce late-phase failures. After considering various modeling options, we chose a multi-state model framework (Putter et al., 2007). Historically, such models have been applied in retrospective analyses and predictions (de Wreede et al., 2010; Beyer et al., 2020; Kunzmann, 2023; Krishnan et al., 2021), whereas for trial planning they have generally been limited to late-phase settings with a focus on PFS and OS, often excluding response as it is not a primary endpoint in confirmatory trials (Xia et al., 2016; Erdmann et al., 2023; Ristl et al., 2021).

To bridge this gap, we developed a multi-state model encompassing states such as “stable disease,” “response,” “progression,” “death,” “no further radiological assessments,” and “terminal drop-out” to evaluate it for early-phase trial planning. Physicians are more accustomed to assumptions on response rates, median PFS, and OS times rather than transition hazards, which are rarely published. Thus, we derive calculations for OS and PFS curves and response probabilities under constant transition hazards, facilitating alignment with physician assumptions. This approach allows rapid evaluation of parameter constellations, enabling effective planning before more complex simulation studies are initiated.

Such a more complex simulation study also incorporates practice-relevant aspects like patient recruitment, assessment schedules and varying analyses time points. We compare our multi-state simulation set-up with separate trial planning for each endpoint and evaluate the operational characteristics in terms of correct/ false go/stop probabilities.

Preliminary observations suggest that integrating endpoints through simulation enhances early-phase decision-making, reduces costly trial failures, optimizes resources, and improves oncology drug development success rates.

Keywords:
Multi state models; Objective response; Progression free survival; Overall survival; Oncology.

References:

A. Delgado and A. K. Guddati (2021): Clinical endpoints in oncology - a primer. American journal of cancer research 11, 1121-31.

Putter et al. (2007): Tutorial in biostatistics: competing risks and multi-state models. Stat Med., 26(11):2389-430.

L. C. de Wreede et al. (2010): The mstate package for estimation and prediction in non- and semi-parametric multi-state and competing risks models. Comput Methods Programs Biomed., 99(3):261-74.

U. Beyer et al. (2020): A multistate model for early decision-making in oncology. Biometrical Journal 62, 550-67.

K. Kunzmann (2023): oncomsm: Bayesian Multi-State Models for Early Oncology. https://github.com/Boehringer-Ingelheim/oncomsm.

S. M. Krishnan et al. (2021): Multistate model for pharmacometric analyses of overall survival in HER2-negative breast cancer patients treated with docetaxel. CPT Pharmacometrics Syst Pharmacol., 10:1255–1266.

F. Xia et al. (2016): A Multi-State Model for Designing Clinical Trials for Testing Overall Survival Allowing for Crossover after Progression. Stat Biopharm Res., 8(1):12-21.

A. Erdmann et al. (2023): Oncology clinical trial design planning based on a multistate model that jointly models progression-free and overall survival endpoints. arXiv:2301.10059v2.

R. Ristl et al. (2021): Delayed treatment effects, treatment switching and heterogeneous patient populations: How to design and analyze RCTs in oncology. Pharm Stat, 20: 129–145.

A nonparametric Bayesian approach for high-dimensional causal effect estimation in survival analysis

T. Jacobs[1], W. van Wieringen[1,2], S. van der Pas[1]

[1] Vrije Universiteit Amsterdam, The Netherlands
[2] Amsterdam UMC, The Netherlands

E-mail: t.jacobs2@amsterdamumc.nl

Abstract:
Estimation of causal effects on survival in the presence of many confounders is hampered by the high-dimensionality of the data. Published work often addresses either high-dimensionality or survival data, but rarely both. We contribute to this gap by studying high-dimensional causal inference methods that account for confounding bias and the complexities introduced by censored observations.

We present a nonparametric Bayesian method to estimate causal effects in high-dimensional, right-censored, and interval-censored survival data. Our approach employs a Bayesian ensemble of Additive Regression Trees (BART) with global-local shrinkage priors on the leaf node parameters (i.e., step heights). This contrasts with existing methods that induce sparsity through the tree skeletons. By focusing shrinkage on the leaf parameters, our method retains all covariates in the model. This reduces the risk of omitting relevant confounders and ensures compliance with the unconfoundedness assumption. The ensemble of Bayesian trees also captures complex, nonlinear relationships between covariates, treatment, and survival outcomes.

We use an efficient Reversible Jump Markov Chain Monte Carlo (RJMCMC) algorithm to sample from the posterior distribution of both the regression trees and the causal treatment effect. Our framework supports a wide range of global-local shrinkage priors. We demonstrate the performance of our method using the Horseshoe prior, which adapts to various levels of sparsity. The general implementation accommodates any scale mixture prior, providing a fast and flexible computational approach for high-dimensional data.

We evaluate our method across a diverse set of simulated data settings, both sparse and dense. This evaluation showcases the method's robustness and flexibility. In sparse settings, the method effectively identifies confounders, while in dense settings, it captures intricate interactions between covariates, treatment, and outcomes. We demonstrate the practical utility of our approach on real-life data from lung cancer patients.

Keywords:
High-dimensional data; Causal inference; Shrinkage priors; Regression trees; Censored data.

Do commonly used machine learning implementations allow for IPCW to address censoring? A closer look at scikit-learn

L. Klein[1], G. Grieser[1], H. Stahl[1], A. Jahn[1]

[1] UAS Darmstadt, Germany

E-mail: lukas.klein@h-da.de

Abstract:
In 2016 the central German organ transplantation registry (TxReg) was established (Nagel, 2022) to enhance research in organ transplantation. Despite its potential, the data is rarely used for analyses (Nagel, 2022). One possible reason for its limited use could be data issues (Otto et al., 2024). Related to survival analysis, significant issues are short maximum follow-up time and annual reporting schedules with occasional in-between reportings, limiting the ability to create long-term survival predictions.

Potentially motivated by similar issues, predictions based on transplantation registries from other countriesare often derived for specific time points (Nitski et al., 2021; Ershoff et al., 2020; Bhat et al., 2018), using classifiers for probability predictions. When adopting this approach, one common method to address censoring is inverse-probability of censoring weighting (IPCW) (Vock, 2016). With IPCW, instance weights are used in the fitting process to achieve unbiased predictions.

Many machine learning libraries such as scikit-learn accept sample weights. However, the documentation of instance weight arguments for the training process usually lacks the implementation details, where and how weights exactly are used and, finally, if the implementation targets the IPCW approach.

For example, in random forests, weights can be applied during the fitting process, either in the bootstrap sampling or in the split criteria within trees. In algorithms derived from gradient descent, weighted loss functions are used. The weights then also affect the Hessian matrix, influencing the optimization process.

We examined the implementation of commonly used classification methods such as the Random Forest Classifier, the Gradient Boosting Classifier and more to assess if the implementation of sample weights can be used to address censoring by IPCW. To support our findings, we conduct a simulation study following ADEMP (Morris et al., 2019) principles. Our goal was the identification of implementations which are able to successfully produce unbiased predictions, when IPCW is applied with increasing censoring rates. Data for the simulation was generated using a Weibull model with varying censoring rates and normally distributed covariates. The objective was to predict survival at a specific time point, using linear models, dense neural networks, tree-based methods, gradient boosting approaches and a model independent weighted bootstrap approach. For the evaluation, the bias on a separate test dataset was calculated. We also compared model performance on the TxReg data for transplantations from deceased donors.

We provide an overview of different instance weighting implementations in current libraries and demonstrate that there are situations, where some implementations prevent the IPCW approach from being able to correct the bias caused by right censoring.

Keywords:
Machine learning; IPCW; Simulation study; Application; Transplantation.

References:

Deutsche Stiftung Organtransplantation (2024): DSO Jahresbericht 2023. www.dso.de.

N. Wouk (2021): End-Stage Renal Disease: Medical Management. American Family Physician, 104(5).

R. Nagel (2022): Über das Transplantationsregister. transplantations-register.de. https://transplantations-register.de/ueber-das-transplantationsregister.

G. Otto, K. Budde, C. Bara and J. Gottlieb (2024): Das Deutsche Transplantationsregister – eine Analyse der Altdaten 2006–2016. Das Gesundheitswesen. https://doi.org/10.1055/a-2251-5627.

O. Nitski, A. Azhie, F. A. Qazi-Arisar, X. Wang, S. Ma, L. Lilly, K. D. Watt, J. Levitsky, S. K. Asrani, D. S. Lee, B. B. Rubin, M. Bhat and B. Wang (2021): Long-term mortality risk stratification of liver transplant recipients: real-time application of deep learning algorithms on longitudinal data. The Lancet Digital Health, 3(5), e295–e305. https://doi.org/10.1016/S2589-7500(21)00040-6.

B. D. Ershoff, C. K. Lee, C. L. Wray, V. G. Agopian, G. Urban, P. Baldi and M. Cannesson (2020): Training and Validation of Deep Neural Networks for the Prediction of 90-Day Post-Liver Transplant Mortality Using UNOS Registry Data. Transplantation Proceedings, 52(1), 246–258. https://doi.org/10.1016/J.TRANSPROCEED.2019.10.019.

V. Bhat, M. Tazari, K. D. Watt and M. Bhat (2018): New-Onset Diabetes and Preexisting Diabetes Are Associated With Comparable Reduction in Long-Term Survival After Liver Transplant: A Machine Learning Approach. Mayo Clinic Proceedings, 93(12), 1794–1802. https://doi.org/10.1016/j.mayocp.2018.06.020.

D. M. Vock, J. Wolfson, S. Bandyopadhyay, G. Adomavicius, P. E. Johnson, G. Vazquez-Benitez and P. J. O'Connor (2016): Adapting machine learning techniques to censored time-to-event health record data A general-purpose approach using inverse probability of censoring weighting. Journal of Biomedical Informatics, 61, 119–131. https://doi.org/10.1016/j.jbi.2016.03.009.

T. P. Morris, I. R. White and M. J. Crowther (2019): Using simulation studies to evaluate statistical methods. Statistics in Medicine, 38(11), 2074–2102. https://doi.org/10.1002/sim.8086.

Life expectancies and blood-based biomarkers for Alzheimer's disease in primary care

L. Kleineidam[1,2], P. Martino-Adami[3], S. Oezdemir[2], M. Wagner[1,2], A. Ramirez[2,3], A. Schneider[1,2], for the AgeCoDe study group[1]

[1] Department of Old Age Psychiatry and Cognitive Disorders, University Hospital Bonn, Germany
[2] German Center for Neurodegenerative Diseases (DZNE), Germany
[3] Division of Neurogenetics and Molecular Psychiatry Department of Psychiatry, University of Cologne, Medical Faculty, Germany

E-mail: Luca.Kleineidam@ukbonn.de

Abstract:
The timely detection of Alzheimer's disease and the prevention of dementia are public health challenges. Measuring plasma pTau217 in primary care holds promise for successful early detection, but an easily interpretable illustration of its prognostic value for dementia incidence is lacking.

Therefore, we estimated remaining life expectancies (LEs) with and without dementia depending on age, sex, and plasma pTau217 levels (low, intermediate, or high) in the AgeCoDe cohort, including dementia-free primary care patients collected randomly from general practitioner registries in six German cities. We used 1451 individuals with available pTau217 data (mean age=83.7 years). The incidence of dementia and death was assessed during up to eight years of follow-up.

We fitted continuous-time multi-state Markov models (R-package msm) modeling transitions from the healthy state to dementia and from both these states to death. We used age as the time scale. Since dementia and death rates grow exponentially with age, we included age as a time-varying covariate, resulting in a definition of hazards of a Gompertz distribution. Sex and pTau217 were modeled as additional covariates. LEs were estimated using the R-package elect.

As expected, higher age and male sex were associated with shorter LEs with and without dementia. Point estimates for total LEs closely resembled LEs from national registry data.

Intermediate (HR[95%-CI]=3.72 [2.62-5.30]) and high (HR[95%-CI]=6.26 [4.63-8.47]) levels of pTau217 were associated with a higher dementia risk compared to low levels, beyond age and sex, which was reflected in estimated LEs with and without dementia.

Across all ages (80-95 years) and sexes, LEs without dementia were the highest in individuals with low pTau217 levels (Est[95%-CI]=2.4 [2.1-3.1] to 10.7 [10.1-11.2]), lower in individuals with intermediate levels (1.9 [1.5-2.4] to 7.4 [6.8-8.2] years) and the lowest in individuals with high levels (1.6 [1.4-1.8] to 5.9 [5.4-6.4]). LEs with dementia were very low in individuals with low pTau217 levels (0.1 [0.1-0.2] to 0.7 [0.6-1.0] years, 4% [2%-5%] to 9% [5%-14%] of remaining lifetime). However, LEs with dementia were considerably higher in individuals with intermediate levels (0.5 [0.3-0.8] to 2.4 [1.9-3.0] years, 15% [12%-21%] to 33% [25%-41%] of remaining lifetime) and the highest in individuals with high pTau217 (0.7 [0.5-0.9] to 3.1 [2.7-3.6] years, 23% [19%-28%] to 45% [36%-56%] of remaining lifetime).

In conclusion, plasma pTau217 is related to marked differences in expected years lived with and without dementia. Estimating LEs provides an intuitively understandable metric for illustrating the influence of risk factors.

Keywords:
Multistate models; Prognosis; Screening.

References:

P. V. Martino-Adami, M. Chatterjee, L. Kleineidam, S. Weyerer, H. Bickel, B. Wiese, A. Ramirez et al. (2024): Prognostic value of Alzheimer's disease plasma biomarkers in the oldest-old: a prospective primary care-based study. The Lancet Regional Health–Europe, 45.

L. Kleineidam, M. Wagner, J. Guski, S. Wolfsgruber, L. Miebach, H. Bickel, K. Heser et al. (2023): Disentangling the relationship of subjective cognitive decline and depressive symptoms in the development of cognitive decline and dementia. Alzheimer's & Dementia, 19(5), 2056-2068.

C. Bellenguez, F. Küçükali, I. E. Jansen, L. Kleineidam, S. Moreno-Grau, N. Amin, O. Goldhardt et al. (2022): New insights into the genetic etiology of Alzheimer's disease and related dementias. Nature genetics, 54(4), 412-436.

Testing the similarity of healthcare pathways based on transition probabilities - A new bootstrap procedure

Z. K. Lange[1], H. Dette[1], M. Farhadizadeh[2], N. Binder[2]

[1] Ruhr-Universität Bochum, Germany
[2] Universität Freiburg, Germany

E-mail: zoe.lange@ruhr-uni-bochum.de

Abstract:

Background: Establishing a common standard of care across clinics and finding the best treatment strategies for diseases are important goals in the healthcare system. To achieve these goals we work with healthcare pathways of patients, consisting of sequences of treatments and other events like hospital readmissions or diagnosis procedures over time. Working with healthcare pathway data is attractive since this data is collected by clinics routinely and therefore, has a high availability. However, the healthcare pathways of different patients tend to be highly heterogeneous, especially for rare diseases lacking a standardized treatment strategy. With the similarity testing approach, presented here, we can find patterns, namely typical pathways, in this heterogeneous data.

Methods: We model the treatment strategies for two different groups of patients by multistate models X_1 and X_2 and analyze their similarity based on their transition probabilities. That is, we test the hypotheses H_0: $d(P_1,P_2) \geq \epsilon$ versus H_1: $d(P_1,P_2) < \epsilon$. Here, P_1 and P_2 are the transition probability matrices of X_1 and X_2, respectively, while $\epsilon > 0$ is a threshold and d a distance on the space of matrices. Groups with similar healthcare pathways, according to the test, are pooled into one group, representing a typical pathway. Based on these pooled data sets, one can perform further estimation tasks like estimating the chances of recovery for a typical treatment. The increased sample size, that results from pooling similar pathways, yields to more accurate statistical inference, especially in small sample settings. This is crucial to reliably identify the best treatment strategy.

Results: We introduce a parametric bootstrap test that is tailored to our similarity hypotheses above. The special attribute of this test consists in the fact that the estimators used for resampling are calculated with respect to a constraint. We proof the validity of this constrained parametric bootstrap test for different measures of similarity d. In the next three month, we will run simulations for the bootstrap test for different sample sizes and analyze the performance of the test on real prostate cancer data from the "Universitätsklinik Freiburg".

Conclusion: Testing the similarity of healthcare pathways to identify best treatment strategies is a new and promising approach that accounts for small sample sizes and draws on easily available data. Additionally, the constrained parametric bootstrap test can easily be adapted to other settings beyond healthcare pathways and can therefore, be of interested in general similarity testing problems.

Keywords:
Similarity testing; Healthcare pathways; Bootstrap.

Prediction stability of survival models

S. Matijevic[1], C. Yau[1]

[1] Big Data Institute, University of Oxford, Oxford, United Kingdom

E-mail: sara.matijevic@linacre.ox.ac.uk

Abstract:
In clinical settings, survival prediction models are essential for estimating patient outcomes over time, such as time to disease progression or mortality risk. Advancements in statistical and machine learning methods have significantly expanded the number of available survival models. However, many of these models lack the stability required for clinical use, as their outputs heavily depend on the specific development data sample, making predictions highly variable if a different dataset is used. This instability can be detrimental to patient health, as model predictions are crucial for guiding treatment options and supporting personalised medicine. Furthermore, significant prediction instability is likely to undermine clinicians' trust in the model, thereby limiting its adoption in clinical practice.

In this study, we evaluated the stability of six widely used survival models: the Cox Proportional Hazards model, Weibull model, Bayesian Weibull model, Random Survival Forest, DeepSurv, and DeSurv. Using synthetic data and a bootstrapping framework, we first assessed model stability across four levels: consistency of population-level mean predictions, stability in the distribution of predictions, robustness within patient subgroups, and reliability of individual predictions. We then concentrated specifically on the fourth level, individual prediction stability, which is arguably the most clinically relevant. To examine this, we used the SUPPORT dataset and evaluated stability through mean absolute prediction error (MAPE), prediction instability plots, calibration instability plots, and MAPE instability plots.

Our findings indicate that classical statistical models, such as the Cox Proportional Hazards (CPH) and Weibull models, provide more stable predictions than machine learning and deep learning survival models. Specifically, the average MAPE for CPH was 0.0105, significantly lower than DeSurv's average MAPE of 0.0908, illustrating greater variability in the deep learning model's predictions. Plots further corroborated these findings, with CPH demonstrating consistently lower prediction and calibration instability. This pattern was also evident in subgroup stability analyses, where CPH maintained lower MAPE values across both balanced and imbalanced subgroups compared to other models.

As the volume of patient data continues to grow, deep learning survival models represent a scalable and efficient approach for analysing large datasets and deriving novel inferences. Stable and reliable models will enable clinicians to make more informed decisions, ultimately improving patient outcomes. More work is needed to improve the prediction stability of deep learning survival models, so they can achieve reliability comparable to their classical statistical counterparts for safe use in clinical settings.

Keywords:
Prediction stability; Deep learning; Bayesian.

Bayesian joint modeling of bivariate longitudinal and time-to-event data: With application of micro and macro vascular complication in people with type 2 diabetes and hypertension

M. Mekonen[1], E. Otranto[1], A. Alibrandi[1]

[1] University of Messina, Italy

E-mail: mequanent2007@gmail.com

Abstract:

Researchers often collect data on chronic disease, including multiple longitudinal measures and time-to-event outcomes. When the time-to-event and the longitudinal outcomes are associated, modeling them separately may give biased estimates. Univariate joint modeling of longitudinal and time-to-event outcomes is an effective method for evaluating their association. However, this model-based analysis can lead to biased estimates when multiple longitudinal outcomes are significantly correlated and deviate from a multivariate normal distribution. We used a bivariate joint model with a skewed multivariate normal distribution to provide a flexible approach for non-symmetrical and correlated longitudinal outcomes. This involved a bivariate linear mixed effects model for longitudinal outcomes and a Cox proportional hazards model for time-to-event outcomes, linked through shared random effects. We estimate the parameters using a Bayesian framework and implement Markov chain Monte Carlo methods via R with JAGS software. The methods are demonstrated using retrospective cohort data from individuals with type 2 diabetes and hypertension at Felgehiwot Referral Hospital in northern Ethiopia. We conduct simulation studies to evaluate the performance of the proposed method. The bivariate linear mixed effects model shows a significant positive relationship between the trajectory of systolic blood pressure and fast blood glucose, which increases significantly over time. The risk of experiencing microvascular complications increased as the subject-specific change rate in fast blood sugar and systolic blood pressure measurements increased (hazard ratio = 1.55, 95% confidence interval: 1.067 to 2.556, and hazard ratio = 3.18, 95% confidence interval: 0.62 to 13.243, respectively). Baseline body mass index (hazard ratio = 1.768, 95% confidence interval: 1.232 to 2.55) and triglycerides (hazard ratio = 1.685, 95% confidence interval: 1.185 to 2.418) were positively associated with the risk of microvascular complications. Our studies suggest a strong and significant positive relationship between the patterns of blood glucose levels and systolic blood pressure. Over time, increases in both blood glucose levels and systolic blood pressure raised the risk of microvascular and macrovascular complications. In bivariate joint modeling, using a skewed multivariate normal distribution instead of a normal distribution makes the model fit better and gives more accurate parameter estimates for the longitudinal biomarker.

Keywords:

Bayesian inference; Bivariate joint model; Longitudinal and time to event data; Type 2 diabetes and hypertension; Micro and macro vascular complications.

References:

M. W. Mekonen, K. A. Birahan , D. M. Chekole and E. A. Derso (2020): Determinants of overall survival of kidney failure for patients receiving dialysis in Saint Geberial General Hospital, Addis Ababa, Ethiopia. Journal of Kidney 6(1), 180, doi-10.35248/2472-1220.20.6.180.

M. W. Mekonen, A. T. Bayew and T. J. Lakew (2022): Factors affecting the duration of gestation among women taking prenatal care at Gondar referral Hospital, Ethiopia. Health Sci Rep., 5(4), e676, doi: 10.1002/hsr2.676.

E. A. Derso, S. Ayalew, A. Eshete and M. Wale (2020): Determinants of time to recovery from obstetric fistula by using the data of University of Gondar Teaching Hospital Fistula Center, Gondar, Ethiopia: A parametric survival regression analysis. Cogent Medicine, 7(1), 1816259, doi: 10.1080/2331205X.2020.1816259.

D. M. Chekole, A. A. Andargie, K. M. Yesuf, M. W. Mekonen, B. M. Geremew and M. Z. Fetene (2022): Multinomial Logistic Regression Analysis on the Prevalence and Determinants of Pneumonia among under five years children using the data of Gondar Referral Hospital. Cogent Public Health, 9(1), doi: 10.1080/2331205X.2022.2029245.

D. A. Mequanente, P. Srinivasan, G. Mallika, P. M. T. Ansari and M. Wale (2022): Time to opportunistic infections and its predictors among HIV-infected children on ART at Gondar University Comprehensive and Specialized Hospital, Ethiopia. Indian Journal of Science and Technology, 15(34), 1675-1682, DOI: 10.17485/IJST/v15i34.1073.

E. Assefa and M. Wale (2022): Determinants of Household Solid Waste Generation Rate in Gondar Town, Northwest Amhara, Ethiopia. International Journal of Statistical Distributions and Applications, 8(3), 55-64, doi: 10.11648/j.ijsd.20220803.13.

A pragmatic approach to the estimation of the interventional absolute risk in continuous time

J. S. Ohlendorff[1], T. A. Gerds[1], A. Munch[1]

[1] University of Copenhagen, Denmark

E-mail: johan.ohlendorff@sund.ku.dk

Abstract:
In medical research, causal effects of treatments that may change over time on an outcome can be defined in the context of an emulated target trial. We are concerned with estimands that are defined as contrasts of the absolute risk that an outcome event occurs before a given time horizon (tau) under prespecified treatment regimens. Most of the existing estimators based on observational data require a projection onto a discretized time scale (van der Laan and Rose, 2011). We consider a recently developed continuous-time approach to causal inference in this setting (Rytgaard et al., 2022) which theoretically allows preservation of the precise event timing on a subject level. Working on a continuous-time scale may improve the predictive accuracy and reduce the loss of information. However, continuous-time extensions of the standard estimators comes at the cost of increased computational burden. In this talk, I will discuss a new sequential regression type estimator for the continuous-time framework which estimates the nuisance parameter models by backtracking through the number of events. This estimator significantly reduces the computational complexity and allows for efficient, single-step targeting using machine learning methods from survival analysis and point processes, enabling robust continuous-time causal effect estimation.

Keywords:
Continuous-time causal inference; Targeted learning; Interventional absolute risk.

References:

S. Rose and M. J. van der Laan (2011): Introduction to TMLE. In: Targeted Learning, Springer Series in Statistics, 67-82, Springer, New York, NY, https://doi.org/10.1007/978-1-4419-9782-1_4.

H. C. Rytgaard, T. A. Gerds and M. J. van der Laan (2022): Continuous-Time Targeted Minimum Loss-Based Estimation of Intervention-Specific Mean Outcomes. The Annals of Statistics 50(5), 2469–2491, https://doi.org/10.1214/21-AOS2114.

Modeling the early-redemption of fixed interest rate mortgages: A survival analysis approach

L. Perotti[1,2], L. A. Grzelak[1,2], C. W. Oosterlee[1]

[1] Utrecht University, Netherlands
[2] Rabobank, Netherlands

E-mail: l.perotti@uu.nl

Abstract:
Accurately modeling the early redemption of fixed-rate mortgages is essential for ensuring fair pricing and effective risk management for financial institutions. Early redemption, or prepayment, often arises from behavioral factors, with relocation to a new house being the predominant driver in the Dutch mortgage market. We address this phenomenon by employing survival analysis techniques to map observable drivers to the distribution of relocation timing. Leveraging a rich dataset of over one million mortgages spanning the past decade, we calibrate our model to capture the underlying dynamics. Furthermore, we analyze the impact of stochastic drivers on the "cost of prepayment" and propose strategies to mitigate this risk. Our findings provide valuable insights into prepayment modeling and offer actionable recommendations for risk management in mortgage portfolios.

Keywords:
Time-to-relocation; Prepayment risk; Survival analysis; Uncertainty quantification; Pricing and hedging.

References:

L. Perotti, L. A. Grzelak and C. W. Oosterlee (submitted 2024): Modeling and Replication of the Prepayment Option of Mortgages including Behavioral Uncertainty. https://arxiv.org/abs/2410.21110.

L. Perotti and L. A. Grzelak (2024): On pricing of discrete Asian and Lookback options under the Heston model. International Journal of Computer Mathematics, 101(8), 889-918, https://doi.org/10.1080/00207160.2024.2363467.

L. Perotti and L. A. Grzelak (2022): Fast sampling from time-integrated bridges using deep learning. Journal of Computational Mathematics and Data Science, 5, 100060, https://doi.org/10.1016/j.jcmds.2022.100060.

Improving Cox regression estimates by using the stochastic approximation expectation-maximization algorithm to handle missing data

E. Peyraud[1,2], J. Jacques[1], G. Metzler[1]

[1] Université Lumière Lyon 2, France
[2] Institut Georges Lopez, France

E-mail: epeyraud@igl-transplantation.com

Abstract:

Introduction: Missing data is a common issue in survival analysis, where traditional approaches to handle missing values, such as single imputation, can often lead to biased results and causing model parameters to tend toward zero. This study explores the Stochastic Approximation Expectation-Maximization (SAEM) algorithm as a more robust alternative, especially for Cox regression. By comparing SAEM to standard imputation techniques, we evaluate its ability to improve parameter estimation across varying levels of missing data.

Methodology: This study adapts the SAEM algorithm to handle missing data in Cox regression, offering a robust approach for incomplete datasets. We assume that data are Missing At Random and that the covariates follow a normal distribution. In our approach, the Cox model includes both observed and missing covariates, with the SAEM algorithm iterating to refine parameter estimates. Rather than the traditional expectation step, SAEM uses a simulation-based approach with Metropolis-Hastings algorithm to generate realistic values for missing data. Each simulated value is evaluated to ensure it matches the conditional distribution of the observed data. Through the stochastic approximation process, SAEM then updates the parameter estimates and log-likelihood with each iteration. The final step updates these parameters, and the algorithm cycle through simulation, approximation, and maximization steps until the estimates stabilize.

Simulation Study: In the simulation study, we generated survival data using a Cox proportional hazards model with five covariates, introducing missing data at varying levels (2%, 5%, 10%, 20%) across the covariates. These levels allowed us to observe how the accuracy of regression coefficient estimates changed as the amount of missing data increased. We compare our method with two well-known imputation methods (Single Imputation by mean, Multiple Imputation by Chained Equation) based on the estimation quality of the model parameters. We show that the proposed method gives a more accurate parameters estimation up to a high rate of missing data.

Perspectives: Future perspectives for this research include testing more complex approaches for the SAEM algorithm, for example by using different assumptions about data. The current approach assumes that there is no correlation between the covariates and we believe that we can increase both performances and stability of the estimation by taking the link between the variables into account. Additionally, applying SAEM to real-world data, particularly liver transplant datasets, where missing covariates are common, would help improve the performance of existing models by better handling missing data while preserving model accuracy and interpretability in clinical settings.

Keywords:
Missing data; Cox regression; Stochastic approximation expectation-maximization.

Assurance methods for designing a clinical trial with a delayed treatment effect

J. Salsbury[1], J. Oakley[1], S. Julious[2], L. Hampson[3]

[1] The School of Mathematics and Statistics, University of Sheffield, United Kingdom
[2] The School of Health and Related Research, University of Sheffield, United Kingdom
[3] Advanced Methodology and Data Science, Novartis Pharma AG, Switzerland

E-mail: jsalsbury1@sheffield.ac.uk

Abstract:
An assurance (probability of success) calculation is a Bayesian alternative to a power calculation. These calculations are becoming more regularly performed in industry, especially in the design of Phase III confirmatory trials. Immuno-oncology (IO) is a rapidly evolving area in the development of anticancer drugs. A common phenomenon that arises from IO trials is one of delayed treatment effects, that is, a delay in the separation of the Kaplan-Meier survival curves. To calculate assurance for a trial in which a delayed treatment effect is likely to be present, uncertainty about key parameters needs to be considered. If uncertainty is not considered, then the number of patients recruited may not be enough to ensure we have adequate statistical power to detect a clinically relevant treatment effect. We present an elicitation technique for when a delayed treatment effect is likely to be present and show how to compute assurance using these elicited prior distributions. We provide an example to illustrate how this could be used in practice.

Keywords:
Assurance; Expert judgement; Prior elicitation; Delayed treatment effects; Probability of success.

References:
J. A. Salsbury, J. E. Oakley, S. A. Julious and L. V. Hampson (2024): Assurance methods for designing a clinical trial with a delayed treatment effect. Statistics in medicine, 43(19), doi:https://doi.org/10.1002/sim.10136.

CORALE project: Cumulative lifetime multi-exposures to ionising radiation and other risk factors and associations with chronic diseases in the CONSTANCES cohort

J. Sauce[1], S. Ancelet[1], C. Mandin[1], P. Renaud[2], J.-M. Métivier[2], A. Chaikh[1], E. Blanchardon[1], D. Broggio[1], C. Gréau[1], C. Vignaud[1], M.-O. Bernier[1], E. Cléro[1], M. Goldberg[3], C. Huet[1], S. Le Got[3], E. Lequy-Flahault[3], A. Isambert[1], A. M. S. Henyoh[1], G. Ielsh[1], C. Michel[1], C. Mukakalisa[1], M. Coeuret-Pellicer[3], H. Roy[1], C. Ribet[3], L. Saey[1], M. Zins[2], O. Laurent[1]

[1] Institut de Radioprotection et de Sûreté Nucléaire (IRSN), PSE-SANTE/SESANE/LEPID, PSE-SANTE/SER/UEM, PSE-ENV/SERPEN/BERAP, PSE-SANTE/SDOS/LEDI, PSE-ENV/SERPEN/BERAD, PSE-SANTE/SDOS/LDRIF, PSE-SANTE/SER/BASEP, F-92260 Fontenay-aux-Roses, France

[2] Institut de Radioprotection et de Sûreté Nucléaire (IRSN), PSE-ENV, PSE-ENV/SERPEN/LEREN, F-13115 Saint-Paul-Lez-Durance, France

[3] Université Paris Cité, Université Paris-Saclay, UVSQ, Inserm, Cohortes Epidémiologiques en population, UMS 11, Villejuif, France

E-mail: justine.sauce@irsn.fr

Abstract:

Introduction: The exposome is a concept encompassing the totality of human exposures from conception onwards. As a first application of this concept to ionising radiation (IR), the CORALE project aims to reconstruct lifetime IR exposures and analyze their associations with chronic diseases in the French general population CONSTANCES cohort. After estimating IR doses to six organs (brain, breast, thyroid, lungs, colon, prostate) from medical and environmental exposures since birth (1941 for the oldest), we focus on evaluating the combined effects of IR exposures and other factors such as smoking and alcohol, on the risks of cancers and cardiovascular diseases.

Methods: A questionnaire was sent to 76,693 volunteers from the CONSTANCES cohort to collect data on several exposures to IR. Medical exposures were estimated by combining data from the questionnaire and the National Health Data System (SNDS), with a thorough review of the literature on French medical imaging practices. Environmental exposures related to geography were reconstructed using residential histories linked with radioactivity databases, while those based on lifestyle were estimated using exposure frequency and age. To assess the combined effects of these cumulated IR exposures treated as time-dependent variables, along with other risk factors, we will apply advanced survival statistical models, including multivariate regression and Bayesian approaches to identify specific exposure profiles.

Results: We estimated lifetime exposure to IR for 30,964 respondents. The mean age at the end of follow-up in 2020 is 55, with a standard deviation of 13 years. We assessed doses, considering year and/or age at exposure and sex when applicable, from several medical procedures, including diagnostic nuclear medicine and CT scans, mammograms, panoramic dental radiographs and chest X-rays; and from environmental factors, including radon, telluric and cosmic radiation (ground-level cosmic rays and those received during air travel), fallout from Chernobyl accident and atmospheric nuclear tests, and seafood consumption. The average lifetime dose to the colon (often used to study association with solid cancers) is estimated to be 90 mSv. Between 5% and 95% of the population received dose below 27 mSv and 198 mSv, respectively. We identified that 37.4% of respondents consumed alcohol regularly, while 48.7% had a history of smoking.

Discussion: Multi-exposure analyses are essential for understanding complex interactions between environmental and behavioral factors and their combined impact on health. We are developing models to explore the associations of these co-exposures with chronic disease incidence. Further extension to other risk factors, such as chemical, is underway.

Keywords:
Constances cohort; Ionising radiation; Multi-exposure; Chronic diseases; Survival analysis.

Bootstrapping LASSO-type estimators in Cox frailty models

L. Schemet[1]

[1] Universität Augsburg, Germany

E-mail: lena.schemet@uni-a.de

Abstract:
The principle of Occam's Razor states that among several plausible explanations for a phenomenon, the simplest is best. Applied to regression analysis, this implies that the smallest model that fits the data is best. Therefore, in terms of analyzing high-dimensional time-to-event data, variable selection techniques are required, if we want to follow the principle of Occam's Razor. One way to achieve variable selection is the use of LASSO-regularization. However, LASSO-regularization does not allow for easy derivation of confidence intervals or p-values for the estimated coefficients, especially in complicated settings like the Cox Frailty Model. We propose bootstrap-based methods to derive confidence intervals for regularized Cox Frailty Models with time-varying effects, more specifically for Hohberg and Groll's (2024) Coxlasso model. We simulate high-dimensional time-to-event data with random effects and time-varying effects to validate the bootstrap approach empirically. The methods are illustrated using a real data example. Future research directions include the theoretical proof of the empirical research.

Keywords:
Bootstrap; Inference; Regularization; Frailty.

References:
M. Hohberg and A. Groll (2024): A Flexible Adaptive Lasso Cox Frailty Model Based on the Full Likelihood. Biometrical Journal, 66(7), https://doi.org/10.1002/bimj.202300020.

Propagator methods for survival analysis

J. Schlecker[1,2,3], I. Kurth[1,3,4], W. W. Hadiwikarta1[1,3,4]

[1] Deutsches Krebsforschungszentrum Heidelberg, Germany
[2] Department of Physics and Astronomy, Heidelberg University, Germany
[3] National Center for Radiation Research in Oncology (NCRO), Heidelberg Institute for Radiation Oncology (HIRO), Heidelberg, Germany
[4] German Cancer Consortium (DKTK), Core Center Heidelberg, Heidelberg, Germany

E-mail: julian.schlecker@dkfz-heidelberg.de

Abstract:
To address a number of shortcomings of conventional survival analysis methods such as Kaplan-Meier Curves and Cox Regression techniques regarding assumptions and requirements imposed on the survival function and its covariate dependence, we introduce a novel survival model, inspired by and as a synthesis of multistate Markov Chain and diffusive Stochastic Process models. By conceptualising the pa- tient as a particle moving stochastically within a compound state-space of discrete categorical and continuous numerical variables that represent the disease state at any given time, we apply methods from Quantum and Statistical Physics to describe the patient's trajectory in this space. This model introduces randomness as a conceptual result of the dynamics of obscured variables, offering distinct interpretability. We treat this perceived randomness due to lack of information with a transition probability function, the propagator, which assigns probability to every possible progression of states given our knowledge. The Regression parameters in this picture have direct meaning in terms of generalised deterministic and random forces acting on the patient. Moreover, assumptions are made at a patient-near level, enhancing clarity in what is required for applicability. Furthermore, this new model provides a more detailed interpretation of the shape of survival curves, naturally considering interactions between parameters. In the presented model, the survival curve is a resultant quantity of the description, in contrast to methods assuming its shape or dependence on covariates a priori. With the propagator, we describe the statistical evolution of patient status, allocating space in the description to the combined effects of both observables and "unobserv- ables". In the end, we aim to aid treatment strategy decisions in application of this survival model. We present the conceptual framework and construction of this model, em- phasizing how it deals with perceived randomness.

Keywords:
Propagator; Stochastic process; Diffusion; Quantum physics; Novel method.

Multi-state models for individualized treatment response prediction and risk assessment in multiple myeloma

S. Schwick[1], S. More[1], H. Fröhlich[1,2]

[1] Department of Bioinformatics, Fraunhofer Institute for Algorithms and Scientific Computing (SCAI), Sankt Augustin, Germany

[2] Bonn-Aachen International Center for Information Technology (b-it), University of Bonn, Bonn, Germany

E-mail: sebastian.schwick@scai.fraunhofer.de

Abstract:

Multiple Myeloma (MM) is a complex and highly heterogeneous hematologic malignancy. Despite frequent advancements in treatments, MM remains incurable, with patients experiencing multiple relapses at various disease stages and a wide range of overall survival. This heterogeneity in disease progression necessitates the need for a personalized treatment approach. Novel treatments for relapsed/refractory MM (RRMM) such as chimeric antigen receptor (CAR) T-cell immunotherapy have shown promising results but may come with significant risks for adverse events. Therefore, thorough research on eligibility criteria for CAR T-cell therapy and risk assessment is critical. Recent studies have shown the advantage of a multi-state framework for individualized risk prediction in newly diagnosed MM (NDMM) by integrating clinical, genomic and treatment variables. We compare and evaluate different multi-state methods and structures, including treatment steps or level of response as possible states, to predict and simulate patient trajectories across multiple lines of therapy. We incorporate longitudinal measurements of MM specific biomarkers (e.g. M protein), treatment regimen and procedures (e.g. stem cell transplantation) as transition specific covariates to assess their impact on the course of the disease. We explore the advantages of utilizing Deep Learning approaches, like neural ordinary differential equations for multi-state models, to reduce model assumptions and increase flexibility. Furthermore, we aim to extend these models to RRMM by incorporating the patient's extensive treatment history and its influence on future treatment outcomes, including the response to CAR T-cell therapy and the risk of adverse events. The developed models will be compared to other machine learning models such as random survival forests and XGBoost. The models will be designed as part of a virtual twin (VT) for MM patients developed by the CERTAINTY (A CEllulaR ImmunoTherapy VirtuAl Twin for PersonalIzed Cancer TreatmeNT) project. Clinical study data (Multiple Myeloma Research Foundation CoMMpass trial: NCT01454297) as well as real world data provided by members of the CERTAINTY consortium will be used for model development. For this purpose, we also consider federated learning approaches, to allow for continuous model improvement as new data becomes available across multiple healthcare centers, while ensuring patient data privacy.

Keywords:

Multistate models; Machine learning; Oncology; Treatment response prediction; Risk assessment.

Double-truncated and censored corporate lifetimes: Likelihood and identification

F. Sieg[1]

[1] University Rostock, Germany

E-mail: fietesieg@gmail.com

Abstract:
We consider a parametric model for corporate lifetimes with a double truncation and censoring structure. The goal is to estimate the parameter of an exponential distribution, which is truncated by a uniform distribution and predominantly subject to censored observations. A maximum likelihood estimator is derived from a point process-theoretic approximation to the underlying truncation process, and we demonstrate both its consistency and asymptotic normality. The developed estimation method is applied to a set of German corporate data, and its behavior is studied through Monte Carlo simulations.

Keywords:
Parametric estimation; Censoring; Truncation.

Combining machine learning methods for subgroup identification in time-to-event data with approximate Bayesian computation for bias correction

H. Stahl[1], L. Klein[1], G. Grieser[1], A. Jahn[1], H. Götte[2]

[1] UAS Darmstadt, Germany
[2] Merck Healthcare KGaA, Darmstadt, Germany

E-mail: henrik.stahl@h-da.de

Abstract:
Personalized medicine is a crucial aspect in finding effective treatments for patients. In clinical development it is essential to identify subgroups of patients who exhibit a beneficial treatment effect, ideally before moving to confirmatory trials. The identified subgroups could be defined by predictive biomarkers with corresponding cut-off values. However, once biomarkers or corresponding cut-offs are selected in a data-driven manner a selection bias is introduced, i.e. the treatment effect within the selected subgroup is overestimated.

In previous work, the approximate Bayesian computation (ABC) algorithm was utilized to correct for this selection bias (Götte et al., 2020). This approach rather covers a reduced range of potential subgroups that are defined by cut-off values. Machine learning (ML)-based subgroup identification methods allow to cover much more potential subgroups with the downside of even greater bias and less interpretable subgroup definitions. Our goal is to extend the ABC algorithm to correct for selection bias also in these situations. Since our research is motivated by clinical trials in oncology, we will focus on time-to-event data such as overall survival or progression-free survival time.

ABC is a simulation approach that selects simulation runs where some particular statistic calculated from trial data at hand is similar to that calculated from simulated data where the true treatment effects are known. The true treatment effects from the selected simulation runs then define an approximation of their posterior distribution that is used for bias correction. Compared to (Götte et al., 2020) ML methods raise additional questions that makes an extension not straight forward: The higher the complexity of the ML approach is the less comparable are the subgroup definitions between the simulation runs. Therefore, next to bias correction also “overlap with true subgroup”, “rate of correct biomarker inclusion” and “similarity in subgroup size” has to be assessed. Depending on the underlying goal of the ML algorithm there is also a higher or lower inherent tendency for bias and a methods potential for correcting that bias needs to be traded off against its potential to identify the “right” patients.

All those aspects are investigated in simulation studies based on the ADEMP framework (Morris et al., 2019). We start with two approaches: model-based partitioning (MOB) (Zeileis et al., 2008; Seibold et al., 2016; Sun et al., 2022) as an ML approach and use LASSO regression (Lipkovich et al., 2017) with treatment interactions as a comparator. In both approaches ABC is investigated for correcting selection bias.

Keywords:
Subgroup Identification; ABC; Bias correction; Machine learning.

References:

H. Götte, M. Kirchner and M. Kieser (2020): Adjustment for exploratory cutoff selection in randomized clinical trials with survival endpoint. Biometrical Journal, 62(3), 627–642, doi: 10.1002/bimj.201800302.

I. Lipkovich, A. Dmitrienko and R. B. (2017): Tutorial in biostatistics: data-driven subgroup identification and analysis in clinical trials. Stat Med, 36(1), 136–196, doi: 10.1002/sim.7064.

T. Morris, I. White and M. Crowther (2019): Using simulation studies to evaluate statistical methods. Statistics in Medicine, 38(11), 2074-2102.

A. Zeileis, T. Hothorn and K. Hornik (2008): Model-based recursive partitioning. Journal of Computational and Graphical Statistics, 17(2), 492–514, doi: 10.1198/106186008X319331.

H. Seibold, A. Zeileis and T. Hothorn (2016): Model-Based Recursive Partitioning for Subgroup Analyses. International Journal of Biostatistics, 12(1), 45–63, doi: 10.1515/ijb-2015-0032.

S. Sun et al. (2022): Comparing algorithms for characterizing treatment effect heterogeneity in randomized trials. Biometrical Journal, 66(1), 2100337, doi: 10.1002/bimj.202100337.

Machine learning for survival analysis: Predicting time-to-event through decomposition

L. Štěpánek[1,2]

[1] Institute of Biophysics and Informatics, First Faculty of Medicine, Charles University, Prague, Czech Republic
[2] Department of Statistics and Probability, Faculty of Informatics and Statistics, Prague University of Economics and Business, Prague, Czech Republic

E-mail: lubomir.stepanek@lf1.cuni.cz

Abstract:

Background/Introduction: Predicting the duration of event-free intervals using multiple covariates is a common task in survival analysis, widely applied in biostatistics. Traditional methods, such as the Cox proportional hazards model and its variants, dominate this field. However, their applicability is constrained by strict statistical assumptions, which may only sometimes hold.

Methods: This study addresses the limitations of traditional approaches by introducing a machine-learning framework for predicting event-free intervals through a novel decomposition of the time-to-event variable. This decomposition reformulates the time-to-event variable into two distinct components: (1) a binary variable indicating whether the event occurred, and (2) a continuous variable representing the timing of the event, if it occurred, or the censoring time otherwise. This separation enables the problem to be approached as a supervised learning task: classification for event occurrence, described using the event component, while the time of the event., i.e., the time component, is used as a predictor. A range of machine-learning classification algorithms can be employed for this task, including logistic regression, support vector machines, decision trees, random forests, naïve Bayes classifiers, and neural networks. By leveraging this decomposition, machine-learning models can be applied with fewer assumptions compared to traditional survival analysis methods.

Results: The proposed method was tested using a COVID-19 dataset comprising multiple explanatory covariates and a response variable representing the time until COVID-19 antibody levels dropped below a laboratory-defined cut-off. Machine-learning models for classification were constructed to predict event occurrences at expected time points. Additionally, the Cox proportional hazards model was used to estimate the same intervals for comparison. Performance metrics indicated similar predictive accuracy between the machine-learning approach and the Cox model, demonstrating that the proposed decomposition-based framework is a viable alternative.

Conclusion: The decomposition of the time-to-event variable into event and time components allows machine-learning models to perform survival analysis with fewer statistical assumptions. When applied to a COVID-19 dataset, the machine-learning approach achieved performance comparable to that of the Cox model. This novel method offers an assumption-minimal alternative for predicting event occurrences, broadening the toolbox available for survival analysis.

Keywords:

Prediction in survival analysis; Time-to-event decomposition; Machine learning; Assumption-minimal methods; Cox proportional hazards model.

References:

L. Štěpánek, F. Habarta, I. Malá and L. Marek (2024): Non-parametric comparison of survival functions with censored data: A computational analysis of greedy and Monte Carlo approaches. In Annals of Computer Science and Information Systems (Vol. 39, pp. 725–730), 19th Conference on Computer Science and Intelligence Systems (FedCSIS), IEEE, https://doi.org/10.15439/2024f223.

L. Štěpánek, J. Novák, O. Vít and L. Marek (2024): Governing the Trade-Off Between Time Period Length and Observations' Number with Machine Learning: A Number of Previous Days Needed for Prediction of Future COVID-19 Positives' Count Using Czech Data. In IFMBE Proceedings (pp. 618–626), Springer Nature Switzerland, https://doi.org/10.1007/978-3-031-62520-6_69.

L. Štěpánek, F. Habarta, I. Malá, L. Štěpánek, M. Nakládalová, A. Boriková and L. Marek (2023): Machine Learning at the Service of Survival Analysis: Predictions Using Time-to-Event Decomposition and Classification Applied to a Decrease of Blood Antibodies against COVID-19. In Mathematics (Vol. 11, Issue 4, p. 819), MDPI AG, https://doi.org/10.3390/math11040819.

L. Štěpánek, J. Dlouhá and P. Martinková (2023): Item Difficulty Prediction Using Item Text Features: Comparison of Predictive Performance across Machine-Learning Algorithms. In Mathematics (Vol. 11, Issue 19, p. 4104), MDPI AG, https://doi.org/10.3390/math11194104.

L. Štěpánek, M. Nakládalová, E. Sovová, L. Štěpánek, A. Boriková, M. Sovová, K. Moravcová, J. Ožana, and L. Jelínek (2023): COVID-19 reduces cardiorespiratory fitness even months after a mild to moderate acute phase: a retrospective cohort study. In Infectious Diseases (Vol. 55, Issue 10, pp. 684–693), Informa UK Limited, https://doi.org/10.1080/23744235.2023.2228408.

L. Štěpánek, F. Habarta, I. Mala and L. Marek (2022): A short note on post-hoc testing using random forests algorithm: Principles, asymptotic time complexity analysis, and beyond. In Annals of Computer Science and Information Systems (Vol. 30, pp. 489–497), 17th Conference on Computer Science and Intelligence Systems, IEEE, https://doi.org/10.15439/2022f265.

L. Štěpánek, M. Nakládalová, M. Janošíková, L. Štěpánek, K. Kabrhelová and A. Boriková (2022): Predictors and characteristics of post-acute COVID-19 syndrome in healthcare workers. In Infectious Diseases (Vol. 55, Issue 2, pp. 125–131), Informa UK Limited, https://doi.org/10.1080/23744235.2022.2136750.

L. Štěpánek, F. Habarta, I. Malá and L. Marek (2021): Reducing the First-Type Error Rate of the Log-Rank Test: Asymptotic Time Complexity Analysis of An Optimized Test's Alternative. In Studies in Computational Intelligence (pp. 281–302), Springer International Publishing, https://doi.org/10.1007/978-3-030-82397-9_15.

L. Štěpánek, F. Habarta, I. Malá and L. Marek (2021): A random forest-based approach for survival curves comparing: principles, computational aspects and asymptotic time complexity analysis. In Annals of Computer Science and Information Systems (Vol. 25, pp. 301–311), 16th Conference on Computer Science and Intelligence Systems, IEEE, https://doi.org/10.15439/2021f89.

A machine learning approach for comparing multiple survival curves: Random forests with reduced assumption dependency

L. Štěpánek[1,2]

[1] Institute of Biophysics and Informatics, First Faculty of Medicine, Charles University, Prague, Czech Republic
[2] Department of Statistics and Probability, Faculty of Informatics and Statistics, Prague University of Economics and Business, Prague, Czech Republic

E-mail: lubomir.stepanek@lf1.cuni.cz

Abstract:

Background/Introduction: Traditional methods for comparing survival curves, such as the log-rank test and Cox proportional hazards models, are widely used in survival analysis. However, these methods rely on strict statistical assumptions, which, if violated, can lead to biased results. Addressing this limitation is critical for improving the robustness of survival analysis.

Methods: This study introduces a novel method for comparing multiple survival curves using a random forest algorithm that minimizes dependency on assumptions. Unlike conventional random survival forests that require a covariate structure and predictors, our approach utilizes only time-to-event data. The method generates variables using common estimators, such as the Kaplan-Meier estimator, which are constant within groups but individualized through imputed missing values into each individual's vector of values after an event occurs. The core principle leverages the decision-making capability of random forests, where each tree can classify data into one, two, or more groups based on survival curves. The fraction of trees that classify into multiple groups, thereby rejecting the null hypothesis of no difference in survival curves, correlates with a calculated ϕ-value analogous to a p-value. Using the Poisson-binomial distribution, we demonstrate that the ϕ-value is a consistent and efficient estimator. Additionally, the Le Cam theorem and Chernoff bound are employed to derive an upper bound for the probability that the ϕ-value falls below a predefined threshold α, offering insights into the statistical power of the method. Tree complexity can be managed through pruning, which reduces the risk of first-type error rate but may also decrease statistical power.

Results: Simulations involving pairs and triplets of statistically distinct and similar survival curves compared the proposed method with the log-rank test and Cox model. Findings indicate that higher levels of tree pruning reduce the risk of first-type error rate while demonstrating a trade-off with statistical power.

Conclusion: This study presents a machine learning-based method for comparing multiple survival curves, providing an almost assumption-free alternative to traditional techniques in survival analysis. The proposed method offers adjustable control over the first-type error rates and demonstrates robustness in varied scenarios, though it may exhibit lower statistical power under certain conditions. This approach represents a possible addition to the toolbox for survival analysis, particularly in scenarios where standard assumptions are challenging to meet.

Keywords:

Survival curves comparison; Machine learning; Random forests; Assumption-free methods; Tree pruning vs. first-type error rate vs. statistical power.

References:

L. Štěpánek, F. Habarta, I. Malá and L. Marek (2024): Non-parametric comparison of survival functions with censored data: A computational analysis of greedy and Monte Carlo approaches. In Annals of Computer Science and Information Systems (Vol. 39, pp. 725–730), 19th Conference on Computer Science and Intelligence Systems (FedCSIS), IEEE, https://doi.org/10.15439/2024f223.

L. Štěpánek, J. Novák, O. Vít and L. Marek (2024): Governing the Trade-Off Between Time Period Length and Observations' Number with Machine Learning: A Number of Previous Days Needed for Prediction of Future COVID-19 Positives' Count Using Czech Data. In IFMBE Proceedings (pp. 618–626), Springer Nature Switzerland, https://doi.org/10.1007/978-3-031-62520-6_69.

L. Štěpánek, F. Habarta, I. Malá, L. Štěpánek, M. Nakládalová, A. Boriková and L. Marek (2023): Machine Learning at the Service of Survival Analysis: Predictions Using Time-to-Event Decomposition and Classification Applied to a Decrease of Blood Antibodies against COVID-19. In Mathematics (Vol. 11, Issue 4, p. 819), MDPI AG, https://doi.org/10.3390/math11040819.

L. Štěpánek, J. Dlouhá and P. Martinková (2023): Item Difficulty Prediction Using Item Text Features: Comparison of Predictive Performance across Machine-Learning Algorithms. In Mathematics (Vol. 11, Issue 19, p. 4104), MDPI AG, https://doi.org/10.3390/math11194104.

L. Štěpánek, M. Nakládalová, E. Sovová, L. Štěpánek, A. Boriková, M. Sovová, K. Moravcová, J. Ožana and L. Jelínek (2023): COVID-19 reduces cardiorespiratory fitness even months after a mild to moderate acute phase: a retrospective cohort study. In Infectious Diseases (Vol. 55, Issue 10, pp. 684–693), Informa UK Limited, https://doi.org/10.1080/23744235.2023.2228408.

L. Štěpánek, F. Habarta, I. Mala and L. Marek (2022): A short note on post-hoc testing using random forests algorithm: Principles, asymptotic time complexity analysis, and beyond. In Annals of Computer Science and Information Systems (Vol. 30, pp. 489–497), 17th Conference on Computer Science and Intelligence Systems, IEEE, https://doi.org/10.15439/2022f265.

L. Štěpánek, M. Nakládalová, M. Janošíková, L. Štěpánek, K. Kabrhelová and A. Boriková (2022): Predictors and characteristics of post-acute COVID-19 syndrome in healthcare workers. In Infectious Diseases (Vol. 55, Issue 2, pp. 125–131), Informa UK Limited, https://doi.org/10.1080/23744235.2022.2136750.

L. Štěpánek, F. Habarta, I. Malá and L. Marek (2021): Reducing the First-Type Error Rate of the Log-Rank Test: Asymptotic Time Complexity Analysis of An Optimized Test's Alternative. In Studies in Computational Intelligence (pp. 281–302), Springer International Publishing, https://doi.org/10.1007/978-3-030-82397-9_15.

L. Štěpánek, F. Habarta, I. Malá and L. Marek (2021): A random forest-based approach for survival curves comparing: principles, computational aspects and asymptotic time complexity analysis. In Annals of Computer Science and Information Systems (Vol. 25, pp. 301–311), 16th Conference on Computer Science and Intelligence Systems, IEEE, https://doi.org/10.15439/2021f89.

Increasing flexibility for the meta-analysis of full ROC curves – A copula approach

F. V. Stoye[1], O. Kuss[2], A. Hoyer[1]

[1] Biostatistics and Medical Biometry, Medical School OWL, Bielefeld University, Germany
[2] Institute for Biometrics and Epidemiology, German Diabetes Center, Leibniz Institute for Diabetes Research at Heinrich Heine University Düsseldorf, Germany

E-mail: ferdinand.stoye@uni-bielefeld.de

Abstract:
The development of new statistical methods for the meta-analysis of diagnostic test accuracy (DTA) studies is a vivid field of research, especially with respect to summarizing full receiver operating characteristic (ROC) curves. Most current approaches to this task utilize Gaussian random effects to account for between-study heterogeneity. While Gaussian random effects are generally tractable using established statistical modeling frameworks and therefore a convenient choice, there is no conceptual reason to restrict the dependence structure to be Gaussian. To increase flexibility in the meta-analysis of ROC curves, we substitute Gaussian random effects with copulas, leading to the ability to directly model the dependence between sensitivity and specificity and an increased control over the estimation procedure. While the resulting models are numerically challenging, they lead to much more flexible and modular model structures when compared to Gaussian random effects. Combined with re-arranging the results reported by DTA studies as being bivariate interval-censored time-to-event data and clinically plausible parametric assumptions for the resulting mixtures in the marginal distributions, this leads to a powerful model to estimate summary ROC curves. An additional advantage of using copulas is the ability to provide a closed-form likelihood, enabling the possibility to use general purpose likelihood optimization strategies. In a simulation study, we use Clayton, Galambos, and Joe copulas with Weibull-binomial and Weibull-normal marginal distributions to compare the resulting models to three alternative approaches from the literature. Our copula models are able to create very flexible model fits with high convergence probabilities and perform similarly to competing models. However, they are also numerically unstable, leading to larger variations in bias as well as lower empirical coverages in the simulation compared to alternative models. This behavior gives rise to the need for a more robust estimation procedure for the copulas. We also show the practical applicability of our copula models to data from a meta-analysis for the screening of type 2 diabetes, leading to plausible estimates for summary ROC curves and the area under the curves.

Keywords:
Copula; Diagnostic test accuracy; Meta-analysis; ROC curve.

A pareto-driven ensemble feature selection approach optimizes biomarker discovery in multi-omics pancreatic cancer studies

J. Zobolas[1], A.-M. George[2], A. López[1], S. Fischer[3], M. Becker[3], T. Aittokallio[1]

[1] Department of Cancer Genetics, Institute for Cancer Research, Oslo University Hospital, Oslo, Norway
[2] Department of Informatics, University of Oslo, Oslo, Norway
[3] Department of Statistics, LMU Munich

E-mail: bblodfon@gmail.com

Abstract:
To address the pressing clinical demands of today, it is crucial to implement models that select minimal, cost-effective features. Feature selection in machine learning aims to fulfill this need by identifying the most predictive biomarkers with minimal redundancy. We have developed a multi-omics ensemble feature selection (EFS) approach that identifies the most significant biomarkers for a given cohort of patients. Our approach leverages multiple machine learning algorithms to discover optimal features for classification, regression, and survival analysis tasks.

The EFS method ranks features using voting theory, ensuring that all ensemble model perspectives are considered. The optimal number of selected features is determined through a Pareto-based knee-point identification method, providing a trade-off between sparsity and performance. When applied to multi-omics datasets from pancreatic cancer studies, our approach successfully identifies minimal biomarkers relevant to both the clinical outcome and the underlying biology of the disease. Overall, EFS offers a reliable and clinically valuable tool for biomarker discovery in cancer research.

Keywords:
Survival analysis; Machine learning; Feature selection; Stability; Ensemble learning; Multi-objective pareto optimization.

References:

A. López (2024): Explainable multi-omics deep clustering model reveals an important role of DNA methylation in pancreatic ductal adenocarcinoma. Nordic Machine Intelligence, 4(2), 26-42, https://journals.uio.no/NMI/article/view/10584.

L. Burk, J. Zobolas, B. Bischl, A. Bender, M. N. Wright and R. Sonabend (submitted 2024): A Large-Scale Neutral Comparison Study of Survival Models on Low-Dimensional Data. https://arxiv.org/abs/2406.04098.

Z. Zhao, J. Zobolas, M. Zicknick and T. Aittokallio (2024): Tutorial on survival modeling with applications to omics data. Bioinformatics, 40(3), btae132, https://doi.org/10.1093/bioinformatics/btae132.

Sponsors

We are very grateful for the support of our sponsors:

AbbVie's mission is to discover and deliver innovative medicines and solutions that solve serious health issues today and address the medical challenges of tomorrow. We strive to have a remarkable impact on people's lives across several key therapeutic areas – immunology, oncology, neuroscience, and eye care – and products and services in our Allergan Aesthetics portfolio.
Website: https://www.abbvie.com/

Deutsche Region der
Internationalen Biometrischen Gesellschaft
(IBS-DR)

The German Region of the International Biometric Society (IBS-DR), established in 1953, is dedicated to promoting biometry in research, education, and application by organizing scientific events and supporting working groups. The region is also a member of the Central European Network, collaborating with neighboring countries to strengthen biometry in the Alps and Central Europe. A special thanks goes to the working group Statistics of Stochastic Processes, led by Jun. Prof. Dennis Dobler and Dr. Jan Feifel, for their valuable support.
Website: https://www.biometrische-gesellschaft.de/

Throughout our 125-year history, Roche has grown into one of the world's largest biotech companies, as well as a leading provider of in-vitro diagnostics and a global supplier of transformative innovative solutions across major disease areas. Our commitment to our people, partners, stakeholders and, most importantly, our patients remains as strong as it was on the first day of our journey.
Website: https://www.roche.com/

Founded in 1834, the Royal Statistical Society (RSS) is one of the world's leading organizations advocating for the importance of statistics and data. It is a professional body for statisticians and other data professionals. The RSS has more than 11,000 members across the world. As a charity, it champions the key role of statistics and data in society, and works to ensure that policy formulation and decision making are informed by evidence for the public good.
Website: https://www.rss.org.uk/

Springer is a leading global scientific, technical and medical portfolio, providing researchers in academia, scientific institutions and corporate R&D departments with quality content through innovative information, products and services.
Website: https://www.springer.com