ASA Section on Lifetime Data Science

LiDS Newsletter

Volume 7, Number 2 – July 2022

ISSN 2473-5159

A section associated with the American Statistical Association

LiDS website: `https://community.amstat.org/lids/home/`
Published newsletters are archived under "Newsletters".

* * * * * * * * * * * *

In Brief

2022 Election Results
Chair-Elect 2023: (Tony) Jianguo Sun
Program Chair-Elect 2023: Pamela Ann Shaw

LiDS Sessions at 2022 JSM
Three invited sessions;
Two topic-contributed sessions;
One contributed session;
One contributed poster session

LiDS Webinar and Short Course Series
Survival Models for Spatial Data
Tuesday, October 18, 2022

2023 LiDS Conference
May 31–June 2, 2023 in Raleigh, North Carolina

Software Review
R package {riskRegression}

* * * * * * * * * * * *

Chair's Message

Greetings, LiDS members! Thus far, 2022 has certainly had its ups and downs; I'll focus on the former. Life hasn't completely returned to normal, but we have come a long way relative to last summer or, especially, the summer before. It's an exciting time of year to the statistician, with the Joint Statistical Meetings just weeks away. Moreover, the Lifetime Data Science Conference (more below) is quickly approaching. Both are in-person events!

The LiDS section is going strong! We currently boast more than 600 members, which represents tremendous growth since LiDS only became a full-fledged American Statistical Association section in 2019. Don't forget to renew your LiDS section membership when you renew your ASA membership, and please spread the word about LiDS to your colleagues.

The year 2022 has featured two LiDS webinars. Dimitris Rizopoulos presented a short-course in March 2022 on "Joint Modeling of Longitudinal and Survival Outcomes". In June, Lu Mao gave a short-course on the "Analysis of Composite Time-to-Event Outcomes". Continuing the momentum of the webinar series, these two events were very well attended by the LiDS membership, totaling over 200 participants. Big thank you's to Professors Rizopoulos and Mao for their excellent courses. In terms of future webinars, LiDS was fortunate to schedule what will be two more excellent events. First, Benjamin Taylor (Lancaster University) is scheduled to give a short-course on "Survival Models for Spatial Data" on October 18, 2022. Then, our first event for winter 2023 will feature Linda Valerie (Columbia University), presenting on "Causal Mediation Analysis of Survival Data". We thank Professors Taylor and Valerie for their support of LiDS and look forward to their courses. The LiDS webinar series remains a prominent event, playing a major role in keeping the ever-increasing LiDS membership connected. We are grateful to Esra Kurum for her outstanding efforts to put together such a stimulating webinar series for 2022–23. It's rewarding that these events have been so well-received by the LiDS members.

The Joint Statistical Meetings are just a few weeks away and LiDS has sponsored three invited sessions and two topic-contributed sessions. One of the latter is dedicated to presentations by the five winners of the 2022 LiDS Student Paper Competition. LiDS thanks Program Chair Jing Ning and Past Program Chair Haoda Fu for putting together the JSM LiDS program and to the members of the LiDS Student Award Committee (Zhezhen Jin, Haoda Fu, Wenjie Wang) for their hard work in reviewing the many submissions and selecting winners.

Preparation is well underway for the 2023 LiDS Conference, to be held May 31–June 2, 2023 at the Marriott City Center in Raleigh, North Carolina. The theme of the conference is "Making an impact in the data science era". The format for 2023 will be similar to that of the previous conferences held at the University of Pittsburgh (2019) and the University of Connecticut (2017). Day 1 (May 31) will consist of short courses by Michael Crowther ("Multi-State Models and other Methods for Complex Data Structures"), Xiaofei Wang and Shu Yang ("Statistical Methods for Time-to-Event Data from Multiple Sources: A Causal Inference Perspective"), as well as former LiDS Chair Jianwen Cai ("Modern Methods and Analysis of Recurrent Event Data"). There will be an opening mixer and poster session on

the evening of May 31. Days 2 and 3 (June 1–2) will each begin with keynote talks, followed by parallel sessions, with a conference dinner to be held on June 1. We are thrilled to have former LiDS Chair Mei-Cheng Wang and Per Andersen as our keynote speakers! Planning is well underway for the conference, with Shanshan Zhao and Wenbin Lu hard at work as Co-Chairs of the Local Organizing Committee. There is still room for more invited sessions, so please contact me at `douglas.schaubel@pennmedicine.upenn.edu` if you are interested in submitting an invited session proposal. Preliminary information on the LiDS Conference will be available at `https://community.amstat.org/lids/home/`. You are encouraged to check back regularly as updates, including registration and more detailed program information, will be frequent during the upcoming fall and winter. Thanks to our LiDS Communications Officer, Wenjie Wang, for maintaining the LiDS website.

The LiDS Annual Business Meeting will be held on Monday, August 8, 5:30–6:30 PM, at the Washington Convention Center (Room CC 101). We encourage LiDS members to attend this meeting as it is a great opportunity to learn more about our Section's activities and to connect with the Executive Committee and other LiDS members. This is also the chance to offer feedback directly to the Executive Committee. We look forward to another informative and lively Annual Business Meeting and hope to see you there!

In closing, I want to thank the LiDS officers on our hard-working Executive Committee, as well as the LiDS membership for their enthusiasm. I look forward to connecting with you at JSM!

Douglas Schaubel, Chair 2022

Nominations Invited for Positions on the Executive Committee

Preparations for the annual election of officers for Sections of the ASA will take place this fall, and as Chair of the Nominations Committee, I am pleased to announce that we are seeking nominations for two positions on the LiDS Executive Committee (EC). In Spring 2023 there will be an election for Chair-Elect, and Program Chair-Elect. The candidate elected in 2023 for the position of Chair-Elect will serve as Chair-Elect in 2024, Chair in 2025, and Past-Chair in 2026. The candidate elected in 2023 for Program Chair-Elect will be Program Chair-Elect in 2024, Program Chair in 2025, and Past-Program Chair in 2026. In accordance with the charter of the Section on Lifetime Data Science, the Nominating Committee consists of the Past Chair of LiDS (Mimi Kim), the Past Program Chair (Haoda Fu), and designated member (Edsel Pena, University of South Carolina); the Past Chair of the Executive Committee serves as Chair of the Nominating Committee. The Nominating Committee shall be responsible for nominating a slate of officers for each annual election which is to be submitted to the ASA by mid-November. There is the opportunity for members of the Section on Lifetime Data Science to nominate candidates for open positions. Such nominations must be signed by at least five members of the Section on LiDS. Nominators should ensure that the nominee is (a) a member of the ASA, (b) a member of the Section on Lifetime Data Science, and (c) is willing to stand for the nomination. Nominations should be sent to the Chair of the Nominations Committee (Mimi Kim, `mimi.kim@einsteinmed.edu`) with a copy to the Secretary (Sharon Xie, `sxie@pennmedicine.upenn.edu`). Please make the subject heading of any nominations "Nomination for 2023 Election for LiDS EC". The deadline for nominations is September 15, 2022. The election is coordinated by the Committee on Sections (COS) of the American Statistical Association.

Mimi Kim, Nomination Committee Chair 2022

2022 Election Results

Chair-Elect 2023 Dr. (Tony) Jianguo Sun is one of Curators' Distinguished Professors at the University of Missouri. Dr. Sun received his Ph.D. in Statistics from the University of Waterloo in 1992 and joined the University of Missouri in 1997. Dr. Sun's research interests include AIDS and cancer research, biostatistics, survival analysis, longitudinal data analysis, and chemometrics. Dr. Sun has published three books and over 160 peer-reviewed research papers in various top journals. Dr. Sun is an elected Fellow of the American Statistical Association, Fellow of the Institute of Mathematical Statistics, and elected Member of the International Statistical Institute.

Program Chair-Elect 2023 Dr. Pamela Ann Shaw is an Adjunct Associate Professor of Biostatistics at the University of Pennsylvania. Dr. Shaw joined the faculty at the University of Pennsylvania after seven years at the National Institute of Allergy and Infectious Diseases. Her statistical research interests include methodology to address covariate and outcome measurement error, the evaluation of diagnostic tests, and design of medical studies. She has a particular interest in the use of biomarker studies to calibrate exposure measurements in the fields of nutritional and physical activity epidemiology and environmental health. More recently she has developed a novel rank test to evaluate a composite survival outcome in the presence of interval censoring. She has continued collaborations in a variety of epidemiologic and clinical studies, with a focus on infectious and chronic disease.

Member Awards: New IMS/ASA Fellows

One out of the 48 newly elected 2022 ASA fellows and three out of the 40 newly elected 2022 IMS fellows are members of our section. Congratulations!

New ASA Fellows

Wei-Ting Hwang, University of Pennsylvania

For statistical leadership of translational and clinical cancer research with major impact on cancer treatment; for methods for biomarker evaluation and risk stratification; and for important contributions to assessment of environmental cancer risks.

New IMS Fellows

Haoda Fu, Eli Lilly and Company

For broad scientific contributions to statistics through machine learning, personalized medicine, survival analysis, and artificial intelligence; For strong leadership in applications of statistics in medical field; and for contributions to the profession.

Zhezhen Jin, Columbia University

For ground-breaking contributions to important areas of statistics and biostatistics, including survival analysis, resampling methods, smoothing methods and statistical computing, and for dedicated and leadership service to the statistics profession.

Jun Yan, University of Connecticut

For significant contributions to survival analysis and copula modeling; for high-impact technology transfer via statistical computing and software development; for influential applications in environmental sciences and public health; and for dedicated student mentoring and service to the profession.

LiDS Student Paper Competition Winners

Congratulations to the following winners of the 2022 LiDS student paper competition!

Benny Ren, University of Pennsylvania, "Combining Mixed Effects Hidden Markov Models with Latent Alternating Recurrent Event Processes to Model Diurnal Active-Rest Cycles"

Di Wang, University of Michigan, "Kullback-Leibler-Based Discrete Failure Time Models for Integration of Published Prediction Models with New Time-To-Event Dataset"

Zehan Yang, University of Connecticut, "Optimal Subsampling for Weibull Accelerated Failure Time Models with Massive Survival Data"

Na Bo, University of Pittsburgh, "A Meta-Learner Framework to Estimate Individualized Treatment Effects for Survival Outcomes"

Can Xie, University of Texas Health Science Center at Houston, "A Flexible-Hazards Cure Model"

LiDS Activities at the 2022 JSM

The JSM 2022 will be held in Washington, D.C. during August 6–11, 2022. We appreciate the session organizers and speakers who chose LiDS as either the primary sponsor or co-sponsor. This year, LiDS has sponsored three invited sessions, two topic-contributed sessions, one contributed session and one contributed poster session as the primary sponsor. Please mark your calendar and attend these sessions to support LiDS.

Invited sessions sponsored by LiDS

- Monday, August 8, 8:30–10:20 AM, CC-152A
  New Statistical Methods for Survival Analysis in Complex Biomedical Studies
  Organizer: Qingning Zhou
  Chair: Qingning Zhou
  Speakers: Donglin Zeng, (Tony) Jianguo Sun, Richard Cook, Jianwen Cai
- Tuesday, August 9, 2:00–3:50 PM, CC-152A
  Data integration and information synthesis in survival analysis Organizer: Hong Zhu
  Chair: Hong Zhu
  Speakers: Li Hsu, Ruth Pfeiffer, Jing Qin, Yu Shen
- Wednesday, August 10, 10:30 AM–12:20 PM, CC-150B
  Real-world Survival Data with Multiple Events: Challenges, Opportunities, and Recent Advancements
  Organizer: Ruosha Li
  Chair: Ruosha Li
  Speakers: Limin Peng, Ross Laverne Prentice, Yifei Sun, Ingrid Van Keilegom

Topic-contributed sessions sponsored by LiDS

- Monday, August 8, 10:30 AM–12:20 PM, CC-144C
  Recent Advances in Clustered Time-to-Event Data
  Organizer: Mei-Ling Ting Lee
  Chair: (Tony) Jianguo Sun
  Speakers: David Oakes, Bernard Rosner, Sin-Ho Jung, Samuel C. Anyaso-Samuel
  Discussant: Jong H. Jeong
- Wednesday, August 10, 8:30–10:20 AM, CC-143C
  LiDS Student Paper Award Winners Organizer: Haoda Fu
  Chair: Haoda Fu
  Speakers: Benny Ren, Di Wang, Zehan Yang, Na Bo, Can Xie

Contributed session sponsored by LiDS

- Monday, August 9, 10:30 AM–12:20 PM, CC-204C
  New Methodology Developments in Analyzing Complex Survival Data
  Chair: Jing Zhang
  Speakers: Prabhashi Withana Gamage, Guowei Li, Theresa Devasia, Myeonggyun Lee, Sundarraman Subramanian

Contributed poster session sponsored by LiDS

- Monday, August 8, 10:30 AM–12:20 PM, CC-Hall D
  Presenter: Madeline Abbott

Jing Ning, Program Chair 2022

LiDS Webinar and Short Course Series

- Short course: **Survival Models for Spatial Data**
- Instructor: Benjamin M. Taylor, Lancaster University, UK
- Sponsor: ASA Lifetime Data Science Section
- Date and Time: Tuesday, October 18, 2022, 11:00 AM–1:00 PM ET (This course will be taught via Zoom)

Course Description Statistical methods for the analysis of survival data are not only applicable in the medical context, but also in many other areas of science and engineering. When survival times are spatially-referenced, some evidence of clustering of high or low times might be apparent on a visual inspection of the data. The question naturally arises as to whether these observed spatial survival patterns can be explained by incorporating appropriate covariates into the model or whether, in order to obtain reliable inferences for model parameters of interest, it is necessary to explicitly model the unexplained spatial variation. In this short course, we give an overview of spatial survival analysis: models, inference, and software and cover three applications of these techniques to real-world challenges.

Part 1: Background and Review of Methodology

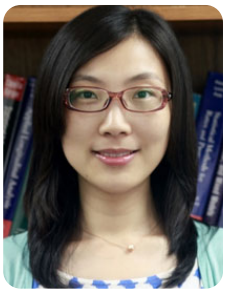
- An introduction to the modelling of spatial stochastic processes
- Combining spatial modelling with survival analyses
- Model selection
- Flexible parametric survival models
- Inference and software for fitting spatial survival models
- Extensions

Part 2: Applications

- Application 1: Modelling Survival in HIV Cohorts in Malawi
- Application 2: Modelling Survival from Colorectal Cancer in Malaysia
- Application 3: Spatial Modelling of Emergency Service Response Times

Esra Kurum, Webinar Committee chair

Call for Invited and Topic-Contributed Session Proposals



The 2023 Joint Statistical Meetings (JSM) will be held in Toronto, Canada during August 5–10, 2023. The theme for 2023 is **"One Community: Informing Decisions and Driving Discovery"**. The LiDS Program Committee is soliciting proposals for an invited session and two topic-contributed sessions for the 2023 JSM. In addition, LiDS will submit up to two proposals for additional invited sessions that are open for competition among the ASA Sections, Interest Groups, the Leadership Support Council, and the Council of Chapters.

An invited session can include two to six participants with a variety of formats. The online submission of invited session proposals opens on July 14, 2022 and closes on September 8, 2022. A topic-contributed session must have five presentations (including discussants). The online submission of topic contributed session proposals opens on November 15, 2022 and closes on December 8, 2022. The regular contributed abstract submission opens on December 1, 2022 and closes on February 1, 2023, and a draft manuscript must be received by May 31, 2023. All proposals should be submitted via the online submission process.

You will need to have the following information to submit your session proposal via the online submission process at `https://ww2.amstat.org/meetings/jsm/2023/submissions.cfm`:

- Session Type (invited)
- Session Subtype (e.g., papers or panel)
- Sponsor: Please select "Lifetime Data Science Section" as the sponsor, and a maximum of two additional potential sponsors can be selected.
- Title of Session
- Brief session description and information for each presenter (title of presentation, name, affiliation, address, phone, email, names of co-authors). Abstracts are NOT required for submission at this time.
- Theme (yes or no): Designate if the session has topics relevant to the JSM theme.
- Applied (yes or no): Designate if the session has topics relevant or will have special appeal to applied statisticians.
- Estimated audience size: Please select the estimated audience size that the session will attract.
- Organizer: name, affiliation, address, phone and email.
- Chair: name, affiliation, address, phone and email.
- Discussants (if any): name, affiliation, address, phone and email.

Please contact Dr. Ying Ding at `yingding@pitt.edu` should you have any question regarding the LiDS-sponsored invited sessions.

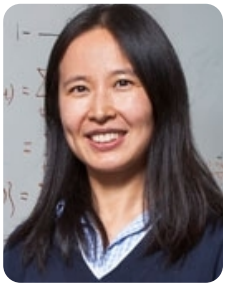
Ying Ding, Program Chair 2023

2023 LiDS Conference

The 2023 LiDS conference will be held May 31–June 2, 2023 in Raleigh, North Carolina at the Marriott City Center. The theme of the conference is "Making an impact in the data science era", retained from the planned 2021 conference which was unfortunately postponed due to the pandemic. Following similar format of the successful conferences at the University of Pittsburgh in 2019 and at the University of Connecticut in 2017, this conference features short courses and an opening mixer/poster session on May 31, two keynote talks and parallel sessions on June 1–2, and a banquet on June 1. We are thrilled to have Dr. Mei-Cheng Wang from the Johns Hopkins University and Dr. Per Andersen from the University of Copenhagen as our keynote speakers. The short courses cover a wide range of topics, including multi-state model, a causal inference approach for time-to-event data in medical research and recurrent events. We will have more than 40 invited scientific sessions on various topics about lifetime data, and we strongly encourage students to present their work in the poster competition. Registration will open later this year. Please mark your calendar and we look forward to meeting you in person at Raleigh!

Shanshan Zhao, Conference Organizing Committee Chair

Treasurer's Report



We received the updated numbers on membership dues, allocations, interests, bank charges and food expense for this year through 5/31/2022. Pending revenue from the short courses in early this year will be added once detailed dollar amount is available. The current balance by May 31, 2022 is $42,232.97.

Beginning Balance	**12/31/2021**	**$43,010.64**
Income		
Membership due and interest		$774.00
Registration		$2,809.15
Net share—CE/Conf/Proceedings		$0.00
Total Income		**$3,583.15**
Expense		
Honorarium		$1,350.00
Meeting overhead		$0.00
Bank/Paypal charges & credit card fees		$10.82
Awards/plaques		$3,000.00
Total Expense		**$4,360.82**
Net Total Income		**($777.67)**
Ending Balance	**5/31/2022**	**$42,232.97**

Yu Cheng, Treasurer 2021–2023

New Articles from Lifetime Data Analysis

Lifetime Data Analysis is the only journal dedicated to statistical methods and applications for lifetime data. The journal advances and promotes statistical science in various applied fields that deal with lifetime data, including actuarial science, economics, engineering, environmental sciences, management, medicine, operations research, public health, and social and behavioral sciences. The journal can be accessed at `https://link.springer.com/journal/10985`.

The April 2022 issue (Volume 28, number 2) of Lifetime Data Analysis has been published:

- A calibrated Bayesian method for the stratified proportional hazards model with missing covariates *by* S. Kim, J.-K. Kim, & K.W. Ahn. Pages 169–193
- Scalable proximal methods for cause-specific hazard modeling with time-varying coefficients *by* W. Wu, J.M.G. Taylor, A.F. Brouwer, L. Luo, J. Kang, H. Jiang, & K. He. Pages 194–218
- Estimation and inference of predictive discrimination for survival outcome risk prediction models *by* R. Li, J. Ning, & Z. Feng. Pages 219–240
- Competing risks regression models with covariates-adjusted censoring weight under the generalized case-cohort design *by* Y. Xu, S. Kim, M.-J. Zhang, D. Couper, & K.W. Ahn. Pages 241–262
- Regression analysis of additive hazards model with sparse longitudinal covariates *by* Z. Sun, H. Cao, & L. Chen. Pages 263–281
- Bayesian penalized Buckley-James method for high dimensional bivariate censored regression models *by* W. Yin, S.D. Zhao, & F. Liang. Pages 282–318
- Bayesian nonparametric dynamic hazard rates in evolutionary life tables *by* L.E. Nieto-Barajas Pages 319–334

Articles in the July 2022 issue (Volume 28, number 3) are:

- A new approach to estimation of the proportional hazards model based on interval-censored data with missing covariates *by* R. Zhou, H. Li, J. Sun, & N. Tang. Pages 335–355
- Mixture survival trees for cancer risk classification *by* B. Jia, D. Zeng, J.J.Z. Liao, G.F. Liu, X. Tan, G. Diao, & J.G. Ibrahim. Pages 356–379
- Longitudinal mediation analysis of time-to-event endpoints in the presence of competing risks *by* T.-T. Vo, H. Davies-Kershaw, R. Hackett, & S. Vansteelandt. Pages 380–400
- Optimum test planning for heterogeneous inverse Gaussian processes *by* C.-Y. Peng, H. Nagatsuka, & Y.-S. Cheng. Pages 401–427
- Semi-supervised approach to event time annotation using longitudinal electronic health records *by* L. Liang, J. Hou, H. Uno, K. Cho, Y. Ma, & T. Cai. Pages 428–491
- Model selection among Dimension-Reduced generalized Cox models *by* M.-Y. Huang, & K.C.G. Chan. Pages 492–511
- Privacy-preserving estimation of an optimal individualized treatment rule: a case study in maximizing time to severe depression-related outcomes *by* E.E.M. Moodie, J. Coulombe, C. Danieli, C. Renoux, & S.M. Shortreed. Pages 512–542

Mei-Ling Ting Lee, Editor-in-Chief, Lifetime Data Analysis

Software Review: {riskRegression}

Scoring risk prediction models

The R-package **riskRegression** provides tools to evaluate and compare the performances of risk prediction models when the outcome is among the following three types (Gerds and Kattan, 2021):

- uncensored binary outcome
- right censored time to event without competing risks
- right censored time to event with competing risks.

In medical applications, risk predictions are made on the probability scale rather than on a classification scale, because the future of a patient is not black or white; it is gray. Also, predictions of the survival time are not appropriate for the patient (Henderson and Keiding, 2005). We thus assume a statistical model that predicts the probability that an event of interest occurs before a given time horizon based on baseline covariates. Such a model is called risk prediction model. Given a list of risk prediction models a key question is which one performs the best. In **riskRegression** prediction performance is measured by the time-horizon dependent Brier score and AUC (Gerds and Kattan, 2021), but not (anymore) by the c-index (Blanche et al., 2019).

During the last years methods were ported from the soon obsolete packages **ModelGood** (uncensored binary outcome), **pec** (right censored outcome) and **timeROC** (right censored outcome) and re-implemented in **riskRegression**. Among the new features that were not available before **riskRegression** we have confidence

intervals for the prediction performance measures and statistical tests to compare prediction models. Packages that implement similar methods are **caret** and **ipred**. At the time of writing the version of **riskRegression** on CRAN is 2022.03.22 and on github 2022.07.05. The example codes below are fully functional from version 2022.07.05 onwards, see https://github.com/tagteam/riskRegression/tree/master/examples/lids-newsletter-riskRegression.R

Scoring risk predictions

We illustrate the riskRegression::Score() function in the setting where the outcome is a right censored time to event variable and there are competing risks. We start by synthesizing the pbc data from the survival package with a recently added function called riskRegression::synthesize(). Then we draw a training data set of size 400.

```
pbc_alike <- synthesize(object=Surv(time,status)~ sex +
    age + logbili + chol + hepato + spiders + protime +
    albumin + platelet + trig + trt + ast,data=pbc)
d <- sim.synth(pbc_alike,n=400,seed=7)
```

We fit two risk prediction models, one based on cause-specific Cox regression (Ozenne et al., 2017) and one based on random survival forest for competing risks (Ishwaran et al., 2014).

```
m1 <- CSC(Hist(time,status)~sex+rcs(age)+rcs(logbili)+rcs
    (protime)+hepato+rcs(chol)+spiders,data=d,fitter="cph
    ")
m2 <- rfsrc(Surv(time,status)~sex+age+logbili+chol+hepato
    +spiders+protime+albumin+platelet+trig+trt+ast,data=d,
    fitter="cph",cause=2,seed=9)
```

Since we have synthesized the real data, we actually have a way to simulate new data from the exact same distribution as the learning data set. We generously provide a validation data set with 3000 observations and call the Score function:

```
nd <- sim.synth(pbc_alike,3000,seed=28)
x <- Score(list("CSC"=m1,"FOREST"=m2),
           data=nd, formula=Hist(time,status)~1,
           cause=2, time=(0:5)*365.25,
           summary="risks",plots=c("calibration","Roc"))
summary(x, times=3*365.25, what="score")$score
```

```
  times      Model          AUC (%)        Brier (%)
1095.75 Null model             50.0 13.1 [12.7;13.6]
1095.75        CSC 73.9 [66.7;81.1] 11.5 [10.8;12.1]
1095.75     FOREST 75.4 [68.7;82.0] 11.2 [10.7;11.6]
```

The confidence intervals are based on estimates of the influence functions of our inverse probability of censoring weighted estimators (Gerds and Schumacher, 2006; Blanche et al., 2013). Note that while the left hand side of the formula specifies the outcome, the right hand side is used to decide which covariates to use when modeling the censoring distribution. Above we have ignored the covariates and estimated the censoring distribution with the so-called reverse Kaplan-Meier method. We are currently working on more flexible censoring models than reverse Kaplan-Meier and Cox regression, and at the same time try to improve the memory usage and computational efficiency of our implementation of the confidence intervals. Stay tuned.

A series of plot functions are applicable to visualize the results of the riskRegression::Score function further. For **ggplot** users there are ggplot2::autoplot versions of the plot functions but these are not yet well documented. Figures 1 and 2 show the result of the following lines:

```
plotRisk(x,times=3*365.25,models=c("FOREST","CSC"))
plotBrier(x)
plotAUC(x) # not shown
plotCalibration(x,times=3*365.25)
plotROC(x,times=3*365.25)
```

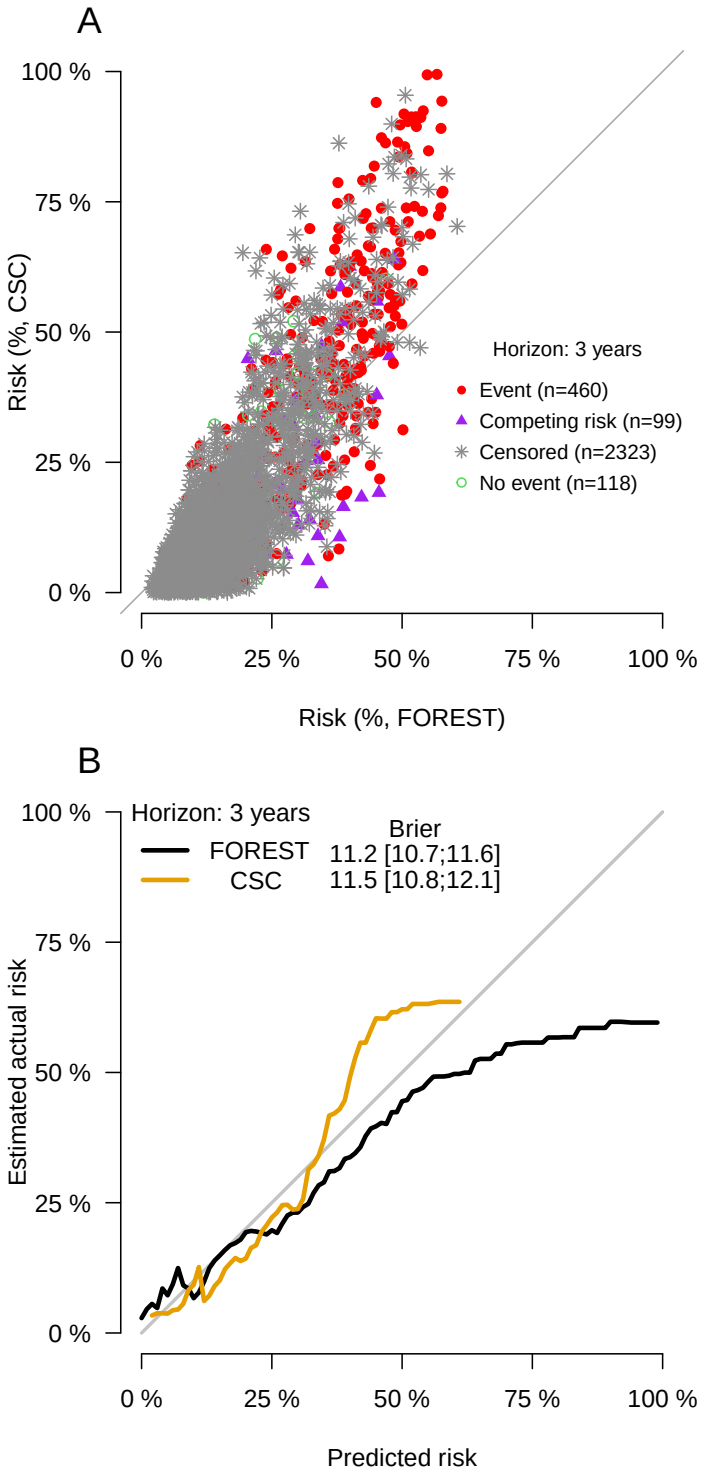


Figure 1: Panel A shows a scatterplot of the predicted risks of the rival models at the 3-year horizon. Panel B shows the calibration plots based on kernel smoothing to estimate the actual risks at the 3-year horizon.

Panel A of Figure 1 shows clearly that even though the overall performances of the two models are not very different, the predicted risks of single patients can be very different (far off the diagonal). It is important to look at the predicted risks.

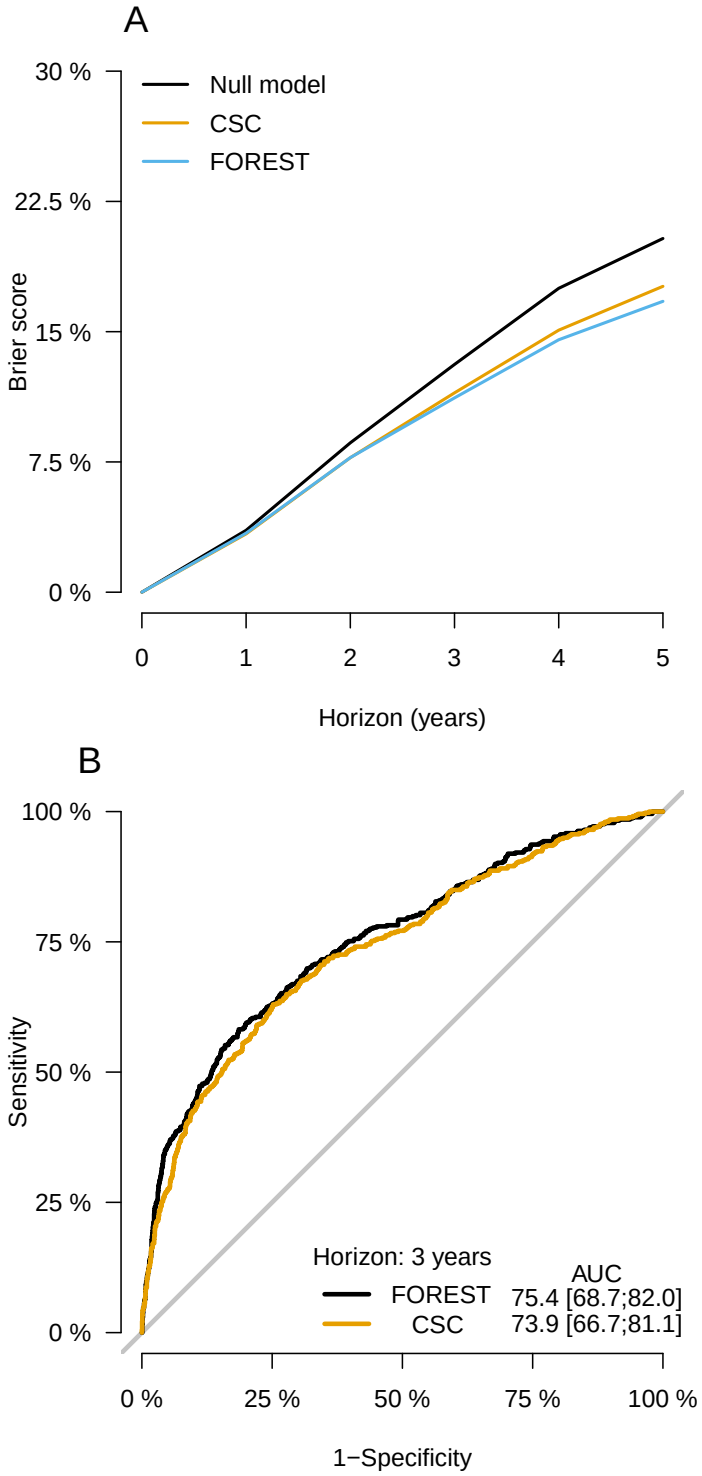


Figure 2: Panel A shows the estimated Brier score across time horizons. Panel B shows the ROC curve at the 3-year horizon.

Writing extensions

For not yet supported risk prediction models, such as the ones provided by your own software, you can either choose to provide a list of matrices which each contain the predicted probabilities at the prediction time horizons, or write a new `S3` method for the generic function `riskRegression::predictRisk()`. However, any cross-validation of a list of risk predictions which involve matrices has to be done by an outer self-made loop. The core of the cross-validation approach of **riskRegression** are the following innocent looking lines of code:

```
model$call$data <- train_set
trained.model <- eval(model$call)
```

In each step of the k-fold or bootstrap loop, the model is fitted again with the current training set. This requires that the model has its `call` saved.

References

Blanche, P., Dartigues, J.-F., and Jacqmin-Gadda, H. (2013), "Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks," *Statistics in Medicine*, 32, 5381–5397.

Blanche, P., Kattan, M. W., and Gerds, T. A. (2019), "The c-index is not proper for the evaluation of-year predicted risks," *Biostatistics*, 20, 347–357.

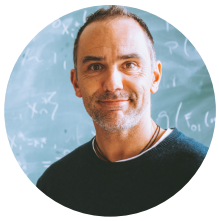
Gerds, T. and Schumacher, M. (2006), "Consistent estimation of the expected Brier score in general survival models with right-censored event times," *Biometrical Journal*, 48, 1029–1040.

Gerds, T. A. and Kattan, M. W. (2021), *Medical Risk prediction models: with ties to machine learning*, CRC Press.

Henderson, R. and Keiding, N. (2005), "Individual survival time prediction using statistical models," *Journal of Medical Ethics*, 31, 703–706.

Ishwaran, H., Gerds, T., Kogalur, U., Moore, R., Gange, S., and Lau, B. (2014), "Random survival forests for competing risks." *Biostatistics*, 15, 757–773.

Ozenne, B., Sørensen, A. L., Scheike, T., Torp-Pedersen, C., and Gerds, T. A. (2017), "riskRegression: Predicting the risk of an event using Cox regression models," *R Journal*, 9, 440–460.



Thomas Alexander Gerds
Professor
Section of Biostatistics
University of Copenhagen
Email: `tag@biostat.ku.dk`

Johan Sebastian Ohlendorff
IT Officer
Section of Biostatistics
University of Copenhagen
Email: `snf991@alumni.ku.dk`