

LiDS Newsletter

Volume 5, Number 2 – July 2020

A section associated with the American Statistical Association

ISSN 2473-5159

LiDS website: <https://community.amstat.org/lids/home>

* * * * *

In Brief

Annual Business Meeting

Friday, August 14, 2020
at 15:30 EDT (virtual)

Nomination for Election 2021

Four positions: Chair-Elect, Secretary, Program Chair-Elect, and Section Representative.

2020 Election Results

Chair-elect: Dr. Douglas Schaubel
Program chair-elect: Dr. Jing Ning
Treasurer: Dr. Yu Cheng

Member Awards

2 IMS fellow and 11 ASA Fellows

LiDS Activities at 2020 JSM

Three invited sessions; five topic-contributed sessions; four contributed sessions; one contributed poster session.

Session Proposal for JSM 2021

Due September 8, 2020 (invited)

Software Review

Recurrent Event Analysis with R packages **reda** and **reReg**

* * * * *

LiDS Officers

Chair:	Nicholas P. Jewell
Chair-Elect:	Mimi Y. Kim
Past Chair:	Jianwen Cai
Secretary 2019-2021:	Joan Hu
Treasurer 2016-2020:	Chiung-Yu Huang
Program Chair (2020):	Zhezhen Jin
Program Chair-Elect :	Haoda Fu
Past Program Chair:	Guoqing Diao
COS Representative:	Xiaonan Xue
Webmaster:	Ker-Ai Lee
Newsletter Editor:	Jun Yan

Chair's Message



Greetings to all our members, sent with the fervent hope that you, and your families, are safe and well in the midst of the COVID-19 pandemic. It has been an extraordinary 6 months to start 2020, and all of our lives are being challenged in significant ways. In addition, and during this pandemic, our country has once again been confronted by its history of racial injustice and inequity. In every component of our lives we must take active commitments to change our practices that are increasingly being shown to have been non-inclusive and discriminatory. We all have a part to play in changing our families, workplaces and institutions in ways that lead

to a fairer and healthier society.

It is important to also recognize the extraordinary efforts being made by many statisticians to confront the challenges of both COVID-19 and racism. We stand together in supporting these significant contributions to our world.

The planning for our third conference on Lifetime Data Science in 2021 is now on pause while we assess the progress of the pandemic and the timing of when opportunities to hold in-person conferences will once again be available. The planning committee is also considering the option of holding an online version as is happening for all other conferences at this moment.

One of these conferences is, of course, the JSM that was originally planned for Philadelphia later this summer. The 2020 JSM has now moved online. The JSM is normally a time to attend many exciting scientific sessions but also provides the opportunity to interact with new and old colleagues from around the world. We hope that you will be able to avail yourself of the online sessions planned by the LiDS section, and consider reaching out to colleagues to ensure that the advantages of meeting in person are not lost by going online. The JSM is also the time when the LiDS section holds its annual business meeting. We are finalizing the timing of this online meeting and will announce details shortly (it will occur between July 31 and August 17, 2020). Please stay posted and join us—your presence will be greatly appreciated.

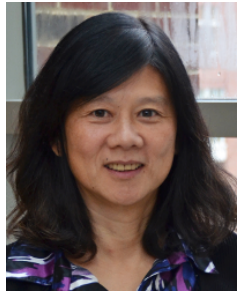
I encourage us all to work together in empathy and kindness, particularly towards our junior colleagues who are experiencing unprecedented obstacles to advancing their careers. This too shall pass and I look forward to greeting many of you in person again soon.

Nicholas P. Jewell, Chair 2020

Nominations Invited for Positions on the Executive Committee

This fall the preparation for the annual election of officers for Sections of the ASA will take place and as Chair of the Nominations Committee I am pleased to announce that we are seeking nominations for four positions on the LiDS Executive Committee (EC). In Spring 2021 there will be an election for Chair-Elect, Secretary, Program Chair-Elect, and Section Representative. The candidate elected in 2021

for the position of Chair-Elect will serve as Chair-Elect in 2022, Chair in 2023, and Past-Chair in 2024. The candidate elected in 2021 for the position of Secretary will serve over the period 2022-2024. The candidate elected in 2021 for Program Chair-Elect will be Program Chair-Elect in 2022, Program Chair in 2023, and Past-Program Chair in 2024. The candidate elected in 2021 for the position of Section Representative will serve over the period 2022-2024.

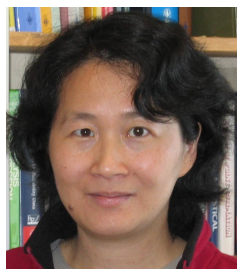


In accordance with the charter of the Section on Lifetime Data Science, the Nominating Committee consists of the Past Chair of LiDS (Jianwen Cai), the Past Program Chair (Guoqing Diao), and one person designated by the Chair (TBD); the Past Chair of the Executive Committee serves as Chair of the Nominating 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 (Jianwen Cai, cai@bios.unc.edu) with a copy to the Secretary (Joan Hu, joan_hu@sfu.ca). Please make the subject heading of any nominations “Nomination for 2021 Election for LiDS EC”. The deadline for nominations is September 15, 2020. The election is coordinated by the Committee on Sections (COS) of the American Statistical Association.

Jianwen Cai, Nomination Committee Chair 2020

Report from the Section Secretary



Becoming the Lifetime Data Science (LiDS) Section of ASA in 2018, our section has now 459 members. The link of our official webpage is <https://community.amstat.org/lids/home>, maintained by Ker-Ai Lee (ka2lee@uwaterloo.ca).

I am delighted to report that 11 out of the 57 newly elected 2020 ASA fellows and two 2020 IMS fellows are either from our section or were members of the Lifetime Data Analysis Interest Group (LiDA-IG). The new ASA Fellows are:

- Kwun Chuen Gary Chan, University of Washington
- Chung-Chou H. Chang, University of Pittsburgh
- Yong Chen, University of Pennsylvania
- Anastasia Ivanova, University of North Carolina – Chapel Hill
- Jialiang Li, National University of Singapore
- Liang Li, MD Anderson Cancer Center
- Sheng Luo, Duke University
- Michael Rosenblum, Johns Hopkins University

- Pamela A. Shaw, University of Pennsylvania
- Rajeshwari Sundaram, National Institute of Child Health and Human Development

The new IMS Fellows are:

- Charmaine Dean, University of Waterloo
- Grace Y. Yi, Western University

Congratulations to all the newly elected IMS and ASA fellows!

The section’s Annual Business Meeting is scheduled for *August 14 at 15:30 EDT*. Look forward to “seeing” you at the meeting.

Joan Hu, Secretary 2019–2021

Treasurer’s Report



The beginning balance of the LiDS Section account on January 1, 2020 was \$42,277.01. The total income during the period between January 1 and May 31 was \$707.86, which included membership dues of \$229.00 and interest \$478.86. There was no expenses during this time period. Then ending balance on May 31, 2020 was \$42,984.87. The ending date was May 31 instead June 30 because the ASA’s finance website has been down for a while.

Beginning Balance	12/31/2019	\$42,277.01
Income:		
Membership (1/1/2020 to 5/31/2020)		\$229.00
Interest (1/1/2020 to 5/31/2020)		\$478.86
Total Income		\$707.86
Expense:		
Total Expense		\$0
Net Total Income		707.86
Ending Balance	5/31/2020	\$42,984.87

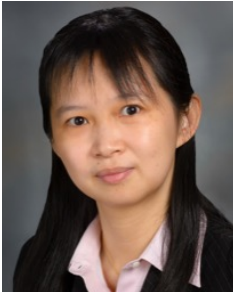
Chiung-Yu Huang, Treasurer 2016–2020

2020 Election Results



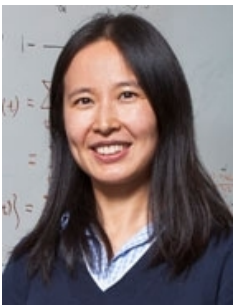
Chair-Elect 2021 Dr. Douglas Schaubel is a Professor at the Department of Biostatistics, Epidemiology and Informatics at the University of Pennsylvania Perelman School of Medicine. Prior to relocating to UPenn, he had spent his entire academic career at the University of Michigan Department of Biostatistics. Dr. Schaubel received his PhD in Biostatistics from the University of North Carolina at Chapel Hill in 2002, under the direction of Dr. Jianwen Cai. His research interests involve the intersection of survival analysis with several areas, including causal inference and facility profiling. His collaborative work has largely been in the areas of end-stage renal disease and chronic liver disease. Dr. Schaubel has published over 160 articles in peer-reviewed

journals. He has held an R01 grant, “Methods for the Analysis of Survival Processes Arising in End-Stage Renal Disease”, which has funded his methodologic research almost continuously since 2005. Dr. Schaubel is an elected Fellow of the American Statistical Association.



Program Chair-Elect 2021 Dr. Jing Ning is an Associate Professor in the Department of Biostatistics at the University of Texas MD Anderson Cancer Center (MDACC). Dr. Ning received a PhD in biostatistics from Johns Hopkins in 2007 and subsequently completed her postdoctoral training at MDACC in 2009. The focus of Dr. Ning’s research has been to develop novel statistical methods and in-

corporate them into better designs and analyses for studies in biomedical research. Her recent methodology contributions include (1) method developments for valid inference by accounting for sampling biases; (2) semiparametric modeling of complex survival data; (3) methods to discover and combine biomarkers for cancer screening, diagnosis, and risk prediction; (4) novel adaptive clinical trial designs; and (5) integrating multiple sources of data to assemble evidence in comparative effectiveness research. She is also strongly committed to providing high-quality statistical collaboration and consultation to basic science and clinical researchers on various applied scientific and medical problems. Her methodological papers appear in *Biometrika*, *Biometrics*, *Biostatistics*, *Journal of the American Statistical Association*, *Journal of the Royal Statistical Society – Series B* and other statistical journals. Her collaborations with biomedical investigators produce publications in high impact medical journals, such as *Annals of Oncology*, *Blood*, *Cancer Discovery*, *Clinical Cancer Research*, *JAMA*, *Journal of Clinical Investigation*, and *Journal of Clinical Oncology*.



Treasurer 2021–2023 Dr. Yu Cheng is Associate Professor of Statistics and Biostatistics at the University of Pittsburgh. She received her PhD in Statistics from the University of Wisconsin-Madison in 2006. Her current methodological research focuses on biomarker evaluation and risk screening, dynamic treatment regimes, association and regression analyses of competing risks data,

quantile association, and causal inference. Prior research experiences include classification, discriminant analysis, and adaptive design. Meanwhile she has participated in various psychiatric studies. Currently she is the PI of an NSF DMS award and a co-investigator on five ongoing NIH-funded grants addressing Alzheimer’s disease, HIV/AIDS, and perinatal weight and impulsive phenotypes. Dr. Cheng has been active in statistical and scientific communities. She has served as interim chair and graduate director of Pitt Statistics Department, President of the ASA Pittsburgh Chapter, a panelist for NIMH and NSF review panels, and Co-chair of local organizing committee for the ASA LiDS conference in 2019.

Member Awards

New IMS Fellow

Charmaine Dean, University of Waterloo: “For her scientifically important contributions to the analysis of count data, disease mapping, spatio-temporal data and more; for her outstanding leadership to the statistical profession, her record of mentorship and for her enormous work in keeping statistics visible at the center of science.”

Yi, Grace Y., University of Western Ontario: “For research excellence in developing theory and methods for the analysis of survival data and longitudinal data in statistical and biostatistical applications, and for world-leading contributions to the analysis of missing and mismeasured data.”

New ASA Fellows

Chan, Gary Hwun Chen, University of Washington: “For outstanding contributions to the methodology of preferential sampling design, observational data and complex lifetime data; for substantive biomedical applications; and for dedication to the mentoring of junior statisticians and health researchers.”

Chang, Chung-Chou H., University of Pittsburgh: “For demonstrated ability to advance the field of biostatistics, being an invaluable collaborator and co-investigator for a large number of clinical researchers, and as a superb teacher and mentor of students throughout the health sciences.”

Chen, Yong, University of Pennsylvania: “For major contributions to the methodology of evidence synthesis; for outstanding research in inference under non-standard conditions, robust inference, and composite likelihood; for bridging statistics and informatics; and for service to the profession.”

Ivanova, Anastasia, University of North Carolina, Chapel Hill: “For outstanding contributions in clinical trial design to advance clinical research and expedite the discovery of treatments that address crucial clinical needs.”

Li, Jialiang, National University of Singapore: “For outstanding contributions to statistical methodology, including diagnostic medicine, non-parametric regression, and personalized medicine; for outstanding mentoring of students; and for sustained research and collaboration involving statistical analysis of medical data.”

Li, Liang, MD Anderson Cancer Center: “For excellent and sustained statistical research and collaboration in the analysis of observational longitudinal cohort studies and chronic disease research, and for outstanding service to the profession.”

Luo, Sheng: Duke University: “For excellent contributions to statistical methods in longitudinal and survival analysis; for scientific contributions to the field of neurological diseases; for exemplary mentorship of students; and for significant service to the profession.”

Rosenblum, Michael, Johns Hopkins University: “For outstanding contributions to statistical methodology and applications, especially with respect to the adaptive design and optimal analysis of randomized trials.”

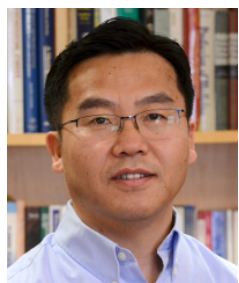
Robert A. Oster, University of Alabama at Birmingham: “For notable contributions to the health science research community in the field of statistics through diverse collaborations; for furthering statistics education in the health sciences; and for service

and exceptional leadership to the ASA and other professional societies.”

Shaw, Pamela, University of Pennsylvania: “For development of innovative statistical methods to address error-prone data, inference in electronic health records, and biomarkers; for outstanding scientific collaborations in nutritional epidemiology, cancer and immunology; and for contributions to statistical education.”

Sundaram, Rajeshwari, National Institute of Child Health and Human Development: “For innovative statistical applications to high-impact problems in reproductive epidemiology and public health; for outstanding methodological contributions to survival analysis and joint modeling of longitudinal and survival data; and for service to the profession.”

LiDS Activities at the 2020 JSM



Due to COVID-19 pandemic, the 2020 Joint Statistical Meetings (JSM) will be held in virtual during August 2-6, 2020

Two Invited Sessions Sponsored by LiDS

Monday, 8/3/2020, 1:00 PM - 2:50 PM; Survival Analysis in Causal Inference Studies; Organizer: Jialiang Li, National

University of Singapore; Chair: Mei-Ling Ting Lee, University of Maryland; Speakers: Mark Van der Laan, Kelly Van Lancker, Torben Martinussen, Jaffer Zaidi, Fei Wan.

Wednesday, 8/5/2020, 10:00 AM - 11:50 AM; Current and Future Challenges in Analyzing Composite Endpoints; Organizer: Li-Shan Huang; Chair: Li-Shan Huang; Speakers: David Oakes, Mei-Cheng Wang, Mei-Ling Ting Lee, Jason Roy.

Two Topic-Contributed Sessions Sponsored by LiDS

Wednesday, 8/5/2020, 1:00 PM - 2:50 PM; Statistical Methods for New Challenges in Lifetime/Complex Data; Organizer: Tony Sit; Chair: Fei Gao; Speakers: Jun Yan, Kin Yau Alex Wong, Yanqing Sun, Tony Sit, Yi Li.

Thursday, 8/6/2020, 3:00 PM - 4:50 PM; Advanced Survival Analysis Tools for Statistical Learning from Complex Scientific Studies; Organizer: Limin Peng; Chair: Wenbin Lu; Speakers: Chiung-Yu Huang, Kwun Chuen Gary Chan, Yu Cheng, Limin Peng.

Four Contributed Sessions Sponsored by LiDS

Monday, 8/3/2020, 10:00 AM - 2:00 PM; Nonparametric and Semiparametric Modeling for Complex Lifetime Data; Chair: Madan Kundu; Speakers: David Zucker, Peng Jin, Sundarraman Subramanian, Anne Eaton.

Tuesday, 8/4/2020, 10:00 AM - 2:00 PM; Statistical Modeling of Lifetime Data: LiDS Section Student Award Session; Chair: Guoqing Diao; Speakers: Teng Fei, Isabelle Weir, Svetlana Eden, Andrew Ying, Shikun Wang, Aliasghar Tarkhan.

Tuesday, 8/4/2020, 10:00 AM - 2:00 PM; Recent Development on the Analysis of Time-to-Event Data; Chair: Christine Mauro; Speakers: Haiyan Yu, Weichi Yao, Qingning Zhou, Xiao Li, Youfei Yu, Yongzhao Shao.

Thursday, 8/6/2020, 10:00 AM - 2:00 PM; Lifetime Risk, Competing Risk, and Recurrent Events; Chair: Yifei Sun; Speakers: Yan Gao, Sarah Conner, Yuxin Zhu, Harrison Reeder, Allison Furgal.

One Contributed Poster Session Sponsored by LiDS

Monday, 8/3/2020, 10:00 AM - 2:00 PM; Presenters: Kristine Gierz, Colin Lewis-Beck, Jung Ae Lee, Koji Kanefuji.

One Invited Session Co-Sponsored by LiDS

Thursday, 8/6/2020, 10:00 AM - 11:50 AM; Various Challenges and Strategies in Analysis of Real-Life Data; Organizer: Joan X. Hu; Chair: Wendy Lou; Speakers: Robert Platt, Wenqing He, Joan X. Hu; Discussant: Rhonda J. Rosychuk.

Three Topic-Contributed Sessions Co-Sponsored by LiDS

Monday, 8/3/2020, 10:00 AM - 11:50 AM; Causal Inference When the Outcome Is Truncated by Death; Organizer: Jessica Gerald Young, Harvard Medical School; Chair: Jessica Gerald Young, Harvard Medical School; Speakers: Mats Stensrud, Zhichao Jiang, Kwonsang Lee, Joseph Hogan.

Tuesday, 8/4/2020, 1:00 PM - 2:50 PM; Estimand Framework and Its Impact on Drug Development in Oncology — Topic Contributed Panel; Organizer: Jonathan Siegel; Chair: Jonathan Siegel; Panelists: Steven Sun, Anja Schiel, Kunthel By, Catherine Njue, Richard J. Cook.

Wednesday, 8/5/2020, 1:00 PM - 2:50 PM; Novel Methods for High-Dimensional and Large-Scale Survival Data; Organizer: Ludovic Trinquart; Chair: Lexin Li; Speakers: Zhe Fei, Gang Li, Shanshan Ding, Jerry Cheng, Simon Bussy.

Zhezhen Jin, Program Chair 2020

Call for Invited and Topic-contributed Session Proposals



The 2021 Joint Statistical Meetings (JSM) will be held in Seattle, Washington during August 7–12, 2021. The LiDS Program Committee is soliciting proposals for an invited session and two topic-contributed sessions for the 2021 JSM.

An invited session can include two to six participants with a variety of formats. The online submission of invited session proposals was open on May 1, 2020 and will close on September 8, 2020. In addition, LiDS can submit up to two proposals for the 24 additional invited sessions that are open for competition among the ASA Sections, Interest Groups, the Leadership Support Council and the Council of Chapters.

A topic-contributed session must have five presentations (including discussants). The online submission of topic contributed session proposals will open on November 12, 2020 and close on December 10, 2020. The regular contributed abstract submission will open on December 3, 2020 and close on February 4, 2021, and a draft manuscript must be received by May 15, 2021. All proposals should be submitted via the online submission process.

The following information is needed:

- Session type (invited)
- Sub-type (e.g., papers or panel)
- Sponsor (ASA Sections, Committee, Journal, Chapter, IMS, ENAR, WNAR, SSC, ICSA, IISA, or Outside Society). Choose from the list of sponsors. Please select “Lifetime Data Science Section “ as the sponsor, and a maximum of 2 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, email)
- Chair (name, affiliation, address, phone, email)
- Discussants (name, affiliation, address, phone, email)

The submission link is at <https://ww2.amstat.org/meetings/jsm/2021/>. July 18 to September 8, 2020 will be the only opportunity to submit invited session proposals.

For further details, please contact 2021 LiDS Program Chair Dr. Haoda Fu (fu_haoda@lilly.com) if you have any questions.

Haoda Fu, Program Chair 2021

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 July 2020 issue (Volume 26, number 3) of Lifetime Data Analysis has been published:

- Semiparametric methods for survival data with measurement error under additive hazards cure rate models *by* Sandip Barui, Grace Y. Yi. Pages 421-450
- Prognostic score matching methods for estimating the average effect of a non-reversible binary time-dependent treatment on the survival function *by* Kevin He, Yun Li, Panduranga S. Rao, Randall S. Sung, Douglas E. Schaubel. Pages 451-470

- A semiparametric additive rates model for the weighted composite endpoint of recurrent and terminal events *by* Xiaowei Sun, Jieli Ding, Liuquan Sun. Pages 471-492
- Bootstrap and permutation rank tests for proportional hazards under right censoring *by* Marc Ditzhaus, Arnold Janssen. Pages 493-517
- Varying coefficient transformation cure models for failure time data *by* Man-Hua Chen, Xingwei Tong. Pages 518-544
- Tree-based modeling of time-varying coefficients in discrete time-to-event models *by* Marie-Therese Puth, Gerhard Tutz, Nils Heim, Eva Münster, Matthias Schmid, Moritz Berger. Pages 545-572
- Multiple event times in the presence of informative censoring: modeling and analysis by copulas *by* Dongdong Li, X. Joan Hu, Mary L. McBride, John J. Spinelli. Pages 573-602
- Group sequential tests for treatment effect on survival and cumulative incidence at a fixed time point *by* Michael J. Martens, Brent R. Logan. Pages 603-623
- Nonparametric estimators of survival function under the mixed case interval-censored model with left truncation *by* Pao-Sheng Shen. Pages 624-637

Articles in the April 2020 issue (Volume 26, number 2) are:

- A new long-term survival model with dispersion induced by discrete frailty *by* Vicente G. Cancho, Márcia A. C. Macera, Adriano K. Suzuki, Francisco Louzada, Katherine E. C. Zavaleta. Pages 221-244
- Assessing the value of a censored surrogate outcome *by* Layla Parast, Lu Tian, Tianxi Cai. Pages 245-265
- Group-sequential logrank methods for trial designs using bivariate non-competing event-time outcomes *by* Tomoyuki Sugimoto, Toshimitsu Hamasaki, Scott R. Evans, Susan Halabi. Pages 266-291
- Penalized full likelihood approach to variable selection for Cox’s regression model under nested case-control sampling *by* Jie-Huei Wang, Chun-Hao Pan, I-Shou Chang, Chao Agnes Hsiung. Pages 292-314
- Semiparametric regression analysis of doubly censored failure time data from cohort studies *by* Shuwei Li, Jianguo Sun, Tian Tian, Xia Cui. Pages 315-338
- Quantile regression-based Bayesian joint modeling analysis of longitudinal-survival data, with application to an AIDS cohort study *by* Hanze Zhang, Yangxin Huang. Pages 339-368
- Parametric and semiparametric estimation methods for survival data under a flexible class of models *by* Wenqing He, Grace Y. Yi. Pages 369-388
- Parametric modelling of prevalent cohort data with uncertainty in the measurement of the initial onset date *by* J. H. McVittie, D. B. Wolfson, D. A. Stephens. Pages 389-401
- A Bayesian approach for semiparametric regression analysis of panel count data *by* Jianhong Wang, Xiaoyan Lin. Pages 402-420

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

Recurrent Event Analysis with R packages `reda` and `reReg`

Recurrent event data arise when the event of interest, such as hospital admissions, infections, or tumor recurrences, can recur in the same individual during follow-up. The standard “time-to-first” event analysis cannot capture the cumulative experience of the recurrent events and could lead to invalid inferences. The R packages `reda` (Wang et al., 2020) and `reReg` (Chiou and Huang, 2020) provide a collection of methods for exploring and analyzing recurrent event data.

Consider a random sample of n subjects and let $N_i(t)$ be the number of events the i th subject experienced over the interval $[0, t]$. Let D be the failure time of interest that could either be a terminal event (e.g., death) or a non-terminal event (e.g., treatment failure). Let C be the potential censoring time for reasons other than the failure event (e.g., study dropouts). The observed data are independent and identically distributed copies $\{N_i(t), Y_i, X_i; t \leq Y_i, i = 1, \dots, n\}$, where $Y_i = \min(D, C)$, $\Delta_i = I(D \leq C)$, $I(\cdot)$ is the indicator function, X_i is a covariate vector, and $N_i(\cdot)$ is observed up Y_i . We illustrate the key features of `reda` and `reReg` with the re-hospitalization data from the `frailtypack` package (Rondeau et al., 2019).

The `Recur()` function prepares the recurrent event data into a `Recur` object used in the packages `reda` and `reReg`. The `Recur` object is an S4 class object that bundles together a set of recurrent times, failure time, and censoring status. The `Recur` object is also used as the formula response for many key functions in `reda` and `reReg`. The following commands create a `Recur` object corresponding to the rehospitalization data:

```
library(reda); library(reReg)
data(readmission, package = "frailtypack")
with(readmission, Recur(t.stop, id, event, death))
```

Error: Subjects having multiple terminal events:
60, 109, 280.

The `Recur()` internally checks if the specified data fits into the recurrent event data framework and detected a possible issue on the data structure. The `show()` method for `Recur` objects presents recurrent events in intervals, where events happened at end of the recurrent episodes with censoring due to (or not) terminal indicated by a trailing + (or *). The following prints the `Recur` object for the first five subjects.

```
with(readmission[1:14,], Recur(t.stop, id, event, death))
```

```
[1] 1: (0, 24], (24, 457], (457, 1037+]
[2] 2: (0, 489], (489, 1182+]
[3] 3: (0, 15], (15, 783*]
[4] 4: (0, 163], (163, 288], ..., (686, 2048+]
[5] 5: (0, 1134], (1134, 1144+]
```

An easy way to glance at recurrent event data is by event plots, which can be created by applying the generic function `plot()` to the `Recur` object when the `reReg` package is loaded. Additionally, the `plotEvents()` function from the `reReg` package allows users to stratify the event plots by discrete variables. The following codes produces event plots with and without stratifying by if the patients received chemotherapy.

```
df0 <- subset(readmission, !(id %in% c(60, 109, 280)))
obj <- with(df0, Recur(t.stop, id, event, death))
plot(obj, legend = "top") # Figure 1
fn <- Recur(t.stop, id, event, death) ~ chemo
plotEvents(fn, data = df0, legend = "top") # Figure 2
```

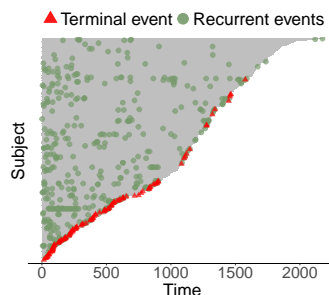


Figure 1: No stratification

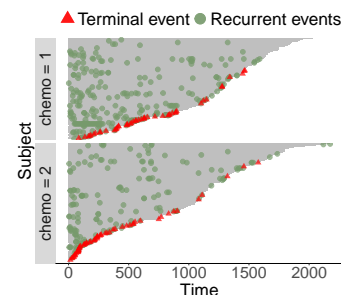


Figure 2: Stratified by chemo.

The mean cumulative function (MCF) is often the focus in the nonparametric analysis of recurrent events. Let $M_i(t) = \mathbb{E}\{N_i(t)\}$ denote the MCF of $N_i(t)$. The Nelson-Aalen estimator (Nelson, 2003) for the MCF takes the form

$$\widehat{M}(t) = \int_0^t \frac{dN(s)}{\delta(s)},$$

where $dN(s) = \sum_{i=1}^k dN_i(s)$, $\delta(s) = \sum_{i=1}^k \delta_i(s)$, $dN_i(s)$ and $\delta_i(s)$ are, respectively, the jump size and at-risk indicator of process i at time s . The MCF can be visualized by plotting the `Recur` object with argument `mcf = TRUE` when the `reReg` package is active, e.g., `plot(obj, mcf = TRUE)`. Alternatively, the `mcf()` function from the `reda` package provides a more sophisticated approach to plot MCFs and make inference. The following example uses the `mcf()` function to visualize MCF estimates stratified by if the patients received chemotherapy.

```
re_mcf <- reda::mcf(fn, data = df0)
plot(re_mcf, conf.int = TRUE, lty = 1:2) +
  ggplot2::theme(legend.position = "bottom") # Figure 3
```

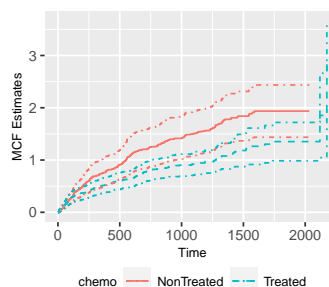


Figure 3: Stratified by chemo

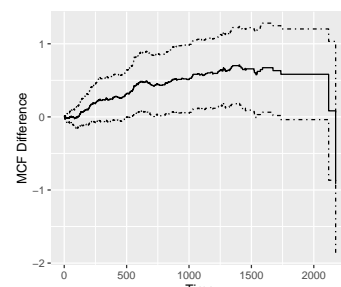


Figure 4: MCF difference

The MCF difference between two groups can be tested via `mcfDiff.test()`, which implements the two-sample pseudo-score tests of (Cook et al., 1996). The following results indicate the MCF estimates are statistically different at a significance level of 0.05. The MCF difference can be plotted with directly by `plot(mcfDiff(re_mcf))`, as shown in Figure 4.

```
reda::mcfDiff.test(re_mcf)
```


Two-Sample Pseudo-Score Tests:

	Statistic	Variance	Chisq	DF	Pr(>Chisq)
Constant Weight	47.49	416.71	5.41	1	0.020 *
Linear Weight	36.56	263.59	5.07	1	0.024 *

Signif. codes:

0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Variance Estimator: robust

The `reReg()` function from the `reReg` package provides methods to fit semiparametric regression models to recurrent event data. A general joint model for the rate function of the recurrent event process and the hazard function of the failure time can be formulated as follow:

$$\lambda(t) = Z\lambda_0(te^{X^T\alpha})e^{X^T\beta}; h(t) = Zh_0(te^{X^T\eta})e^{X^T\theta}, \quad (1)$$

where Z is a latent shared frailty variable to account for association between the two types of outcomes, $\lambda_0(\cdot)$ is the baseline rate function, $h_0(\cdot)$ is the baseline hazard function, and the regression coefficients (α, η) and (β, θ) correspond to the shape and size parameters of the rate function and hazard function, respectively. In contrast to many shared-frailty models that require a parametric assumption, following the idea of Wang et al. (2001), the `reReg()` function implements semiparametric estimation procedures that do not require the knowledge about the frailty distribution. As a result, the dependence between recurrent events and failure event is left unspecified and the proposed implementations accommodate informative censoring.

Model (1) includes several popular semiparametric models as special cases, which can be specified via the `method` argument with the rate function and hazard function separated by “|”. For examples, the joint Cox model of Huang and Wang (2004) is a special case of (1) when $\alpha = \eta = 0$ and can be called by `method = "cox|cox"`; the joint accelerated mean model of Xu et al. (2017) is a special case when $\alpha = \beta$ and $\eta = \theta$ and can be called by `method = "am|am"`. Treating the terminal event as nuisances ($\eta = \theta = 0$), (1) reduces to the generalized scale-change model of Xu et al. (2019), called by `method = "sc|."`. Moreover, users can mix the models depending on the application. For example, `method = "cox|ar"` postulate a Cox proportional model for the recurrent event rate function and an accelerated rate model for the terminal event hazard function ($\alpha = \theta = 0$ in (1)). For inference, the asymptotic variance is estimated from an efficient resampling-based sandwich estimator motivated by Zeng and Lin (2008). The resampling approach is faster than the conventional bootstrap as it only requires evaluating perturbed estimating equations rather than solving them. The following code fits the joint Cox model with 200 (default) resampling replicates.

```
fit <- reReg(fn, df0, method = "cox|cox"); summary(fit)
```

Call:

```
reReg(formula = fn, data = df0, method = "cox|cox")
```

Recurrent event process:

	Estimate	StdErr	z.value	p.value
chemoTreated	-0.189	0.243	-0.779	0.436

Terminal event:

	Estimate	StdErr	z.value	p.value
chemoTreated	0.519	0.311	1.669	0.095

The model shows that patients treated with chemotherapy are hospitalized less often (17% lower) and have a higher risk of death (68% more) than patients not treated with chemotherapy though not significant at the 0.05 level. After a model is fitted, the baseline rate function and hazard function can be visualized by plotting the `reReg()` object. See wenjie-stat.me/reReg/ and www.sychiou.com/reReg/ for the full package documents.

Reference

- Chiou, S. H. and Huang, C.-Y. (2020), *reReg: Recurrent Event Regression*, R package version 1.3.0.
- Cook, R. J., Lawless, J. F., and Nadeau, C. (1996), “Robust tests for treatment comparisons based on recurrent event responses,” *Biometrics*, 52, 557–571.
- Huang, C.-Y. and Wang, M.-C. (2004), “Joint modeling and estimation for recurrent event processes and failure time data,” *Journal of the American Statistical Association*, 99, 1153–1165.
- Nelson, W. B. (2003), *Recurrent Events Data Analysis for Product Repairs, Disease Recurrences, and Other Applications*, Society for Industrial and Applied Mathematics.
- Rondeau, V., Gonzalez, J. R., Mazroui, Y., Mauguen, A., Diakite, A., Laurent, A., Lopez, M., Król, A., and Sofeu, C. L. (2019), *frailtypack: General frailty models: Shared, joint and nested frailty models with prediction; Evaluation of failure-time surrogate endpoints*, R package version 3.0.3.
- Wang, M.-C., Qin, J., and Chiang, C.-T. (2001), “Analyzing recurrent event data with informative censoring,” *Journal of the American Statistical Association*, 96, 1057–1065.
- Wang, W., Fu, H., and Yan, J. (2020), *reda: Recurrent Event Data Analysis*, R package version 0.5.2.
- Xu, G., Chiou, S. H., Huang, C.-Y., Wang, M.-C., and Yan, J. (2017), “Joint scale-change models for recurrent events and failure time,” *Journal of the American Statistical Association*, 112, 794–805.
- Xu, G., Chiou, S. H., Yan, J., Marr, K., and Huang, C.-Y. (2019), “Generalized scale-change models for recurrent event processes under informative censoring,” *Statistica Sinica*, doi:10.5705/ss.202018.0224.
- Zeng, D. and Lin, D. (2008), “Efficient resampling methods for nonsmooth estimating functions,” *Biostatistics*, 9, 355–363.



Sy Han Chiou

Assistant Professor

Department of Mathematical Sciences

University of Texas at Dallas

Email: schiou@utdallas.edu



Wenjie Wang

Research Scientist

Machine Learning, Artificial

Intelligence, and Connected Care

Advanced Analytics and Data Sciences

Eli Lilly and Company

Email: wang.wenjie@lilly.com