9TH STATISTICS & BIOPHARMACY CONFERENCE
19 | 21 SEPT 2022

# Towards Filling the Gaps around Recurrent Events in High-Dimensional Framework: A Systematic Literature Review and Application

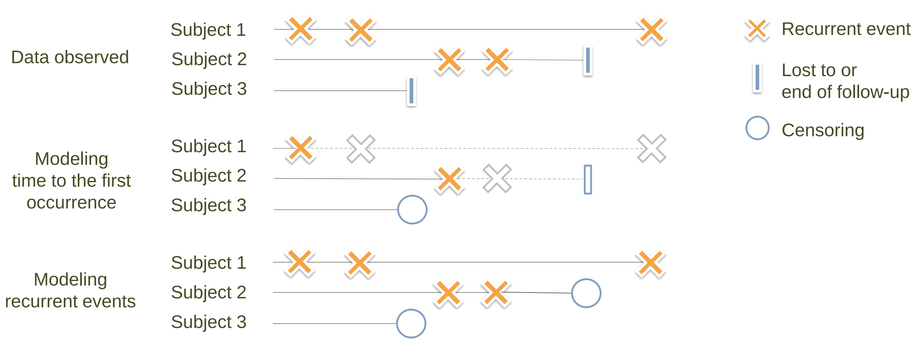
Juliette Murris[1,2,3], Anaïs Charles-Nelson[4,5], Audrey Lavenu[6,7,8], Sandrine Katsahian[1,2,4,5,9]

[1] Inserm, Centre de recherche des Cordeliers, Université de Paris, Sorbonne Université, Paris, France; [2] HeKA, Inria, Paris, France ; [3] RWE & Data, Pierre Fabre, Boulogne-Billancourt, France ; [4] AP-HP, Hôpital Européen Georges-Pompidou, Unité de Recherche Clinique, APHP.Centre, Paris, France ; [5] Inserm, Centre d'Investigation Clinique 1418 (CIC1418) Épidémiologie Clinique, Paris, France ; [6] Université de Rennes 1, Faculté de médecine, Rennes, France ; [7] IRMAR, Institut de Recherche Mathématique de Rennes, Rennes, France ; [8] CIC Inserm CIC 1414, Université de Rennes 1, Rennes, France ; [9] Hôpital Européen Georges Pompidou, Service d'informatique médicale, biostatistiques et santé publique, AP-HP, Paris, France

## CONTEXT & OBJECTIVES

- Study individuals may face **repeated events over time**, such as hospitalizations or cancer relapses (Figure 1)
- In either clinical trials or real-world set, **survival analysis** usually focuses on modeling the time to the first occurrence of the event
- Modern technologies enable data to be generated on **thousands of variables** or observations, as per genomics, medico-administrative databases, disease monitoring by intelligent medical devices
- **Standard statistical models** may no longer be applied when the number of variables studied $p$ is greater than the number of individuals $n$

*Figure 1. Recurrent Event Framework*




### Study objectives

- To identify learning algorithms for analysing/predicting recurrent events in high-dimensional framework
- To apply them along with to standard statistical models in various data simulation settings

## METHODS

### Systematic literature review (SLR)

- Conducted to provide state-of-the-art methodology
- Inclusion criteria – survival analysis in a high-dimensional framework or use of machine learning techniques for recurrent event data
- Exclusion criteria – Bayesian approaches and clinical trial design
- Implication of independent two reviewers to assess publication eligibility

### Statistical analysis

- Standard models applied and extended (Table 1)
- Evaluation criteria included Harrell's Concordance-index[5] (C-index), Kim's C-index[6] and error rate for active variables

*Table 1. Standard Statistical Models for Recurrent Events Analyses*

| Model | Type | Hazard function | Components and specificities |
|---|---|---|---|
| *AG*[1] | Conditional model | $\lambda_i(t) = Y_i(t) \times \lambda_0(t) \times \exp(\beta^t X_i)$ | Recurrent events within individuals are independent and share a common baseline hazard function |
| *PWP*[2] | Conditional model | $\lambda_{ik}(t) = Y_i(t) \times \lambda_{0k}(t) \times \exp(\beta_k^t X_i)$ | Stratified AG, stratum *k* collects all the $k^{th}$ events of the individuals Hazard function for each event |
| *WLW*[3] | Marginal model | $\lambda_{ik}(t) = Y_i(t) \times \lambda_{0k}(t) \times \exp(\beta_k^t X_i)$ | Marginal model, also stratified, calendar time scale and semi-restricted set for subjects at risk Intra-subject dependence |
| *Frailty*[4] | Conditional model | $\lambda_i(t) = Y_i(t) \times \lambda_0(t) \times z_i \times \exp(\beta^t X_i)$ | Random term $z_i$ for each individual to account for unobservable or unmeasured characteristics |

AG = Andersen-Gill; PWP = Prentice, William et Peterson; WLW = Wei-Lin-Weissfeld. $\boldsymbol{X}_i$ a *p*-dimensional vector of covariates, $\beta$ the associated regression coefficients, $\lambda_0(t)$ the baseline hazard function, $Y_i(t)$ an indicator of whether subject *i* is at risk at time *t*

### Data & Simulation scheme

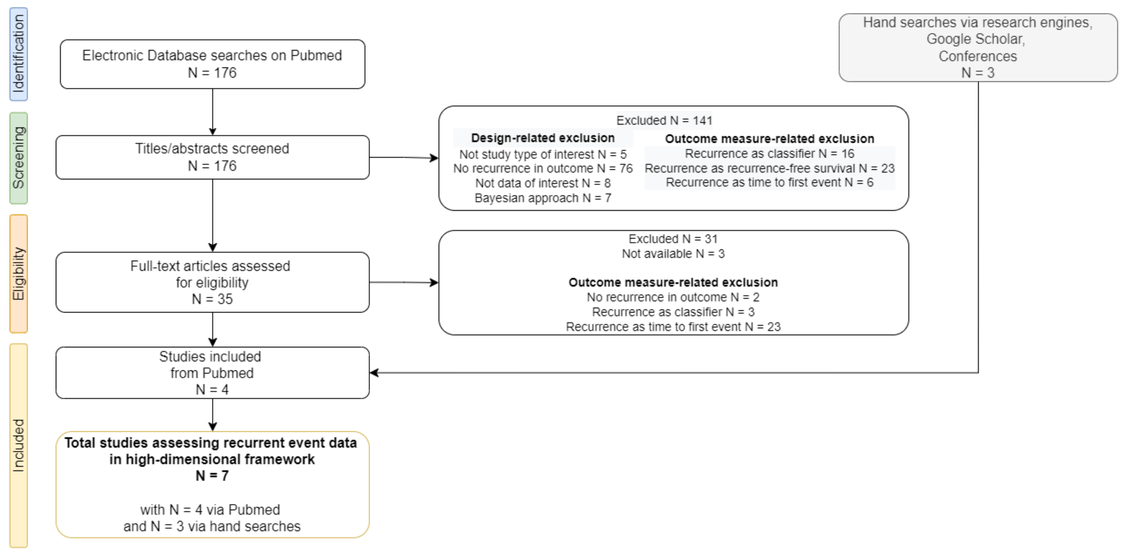
- `simrec` package[7] in R was extended to control for multicollinearity and proportion of active variables (sparse rate)
- 15 scenarios were generated based on the number of subjects ($N = 100$), the censoring rate ($c = 20\%$), the sparse rate ($sp = 0\%, 25\%, 50\%$) and the number of variables ($p = 25, 50, 100, 150, 200$)

## RESULTS

**The SLR, from 176 hits to 7 relevant publications** (Figure 2)

- Recurrence was considered as a classifier (19/176), as a recurrence-free survival outcome (23/176), or as a time-to-first event (29/176)
- 7 publications were identified, consisting in 4 methodological studies, 2 reviews and 1 application paper
- 2 methods with open-sourced code were selected for application: variable selection using **BAR penalty**[8] and **RankDeepSurv** neural network[9]

*Figure 2. Flowchart of included publications in the systematic literature review*

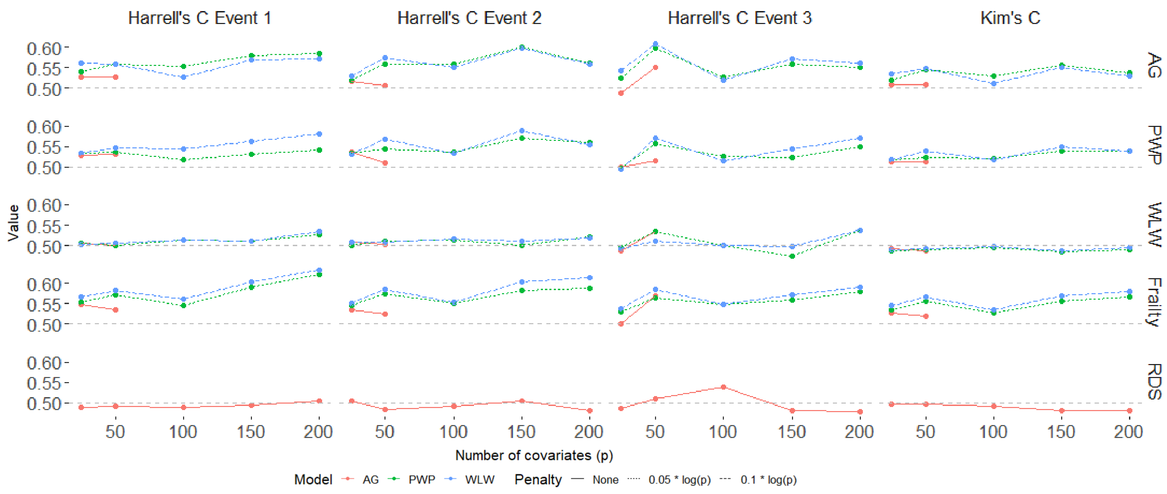



**Take home message 1**
As there are no published guidelines or recommendations, the SLR illustrates authors' caution when dealing with recurrent events and high dimensional data

**Simulation results** (Figure 3)

- As expected,
  - Standard models failed as soon as $p > n$
  - Penalized helped to improve their performance when $p < n$
  - C-indices were around 0.5 when $sp = 0\%$
- Best performance was obtained using penalized frailty model
- Worst performance was observed for WLW and RankDeepSurv
- Kim's C-index was more stable across the different number of covariates and sparse rates

*Figure 3. Impact of the number of covariates on average C-indices with 25% sparse rate*




Average C-indices of the 100 simulated datasets were displayed over the number of covariates. Penalties > 0 were applicable only for standard statistical models, RankDeepSurv was not penalized. Unpenalized standard statistical models did not converge as soon as $p > n$, performance was therefore not available.

## DISCUSSION & CONCLUSION

Standard and identified approaches had not been confronted with one another. This may lead to erratic behavior and confusion when researchers wish to conduct robust and reliable analyses in such a context.

### Study limitations

- More scenarios could be explored and include variations of number of subjects and censoring rate
- Hyperparameters from the BAR penalty method could not be optimized
- Other evaluation metrics could be used e.g., mean square error, mean absolute error, log-likelihood, feature importance

**Take home message 2**
Deep ML approach does not outperform. Besides, there is no scientific consensus on the best performance metric to use.

To the author's knowledge, this work was the first to confront standard methods, variable selection algorithms, and a deep neural network in modeling recurrent events in a high-dimensional framework, and specifically to measure the impact of the number of covariates.

## BIBLIOGRAPHY, ACKNOWLEDGEMENT

1 Andersen PK, Gill RD. Cox's Regression Model for Counting Processes: A Large Sample Study. Ann Stat. 1982;10(4):1100–20.
2 Prentice RL, Williams BJ, Peterson AV. On the regression analysis of multivariate failure time data. Biometrika. 1981;68(2):373–9.
3 Wei LJ, Lin DY, Weissfeld L. Regression Analysis of Multivariate Incomplete Failure Time Data by Modeling Marginal Distributions. J Am Stat Assoc. 1989;84(408):1065–73.
4 Vaupel JW, Manton KG, Stallard E. The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography. 1979 Aug;16(3):439–54.
5 Harrell FE, Lee KL, Mark DB. MULTIVARIABLE PROGNOSTIC MODELS: ISSUES IN DEVELOPING MODELS, EVALUATING ASSUMPTIONS AND ADEQUACY, AND MEASURING AND REDUCING ERRORS. Stat Med. 1996 Feb 29;15(4):361–87.
6 Kim S, Schaubel DE, McCullough KP. A C-index for recurrent event data: Application to hospitalizations among dialysis patients. Biometrics. 2018 Jun;74(2):734–43.
7 Jahn-Eimermacher, Antje, Katharina Ingel, Ann-Kathrin Ozga, Stella Preussler, and Harald Binder. 2015. "Simulating Recurrent Event Data with Hazard Functions Defined on a Total Time Scale." BMC Medical Research Methodology 15 (1): 16.
8 Zhao H, Sun D, Li G, Sun J. Variable selection for recurrent event data with broken adaptive ridge regression. Can J Stat. 2018;46(3):416–28.
9 Jing B, Zhang T, Wang Z, Jin Y, Liu K, Qiu W, et al. A deep survival analysis method based on ranking. Artif Intell Med. 2019 Jul;98:1–9.

JM benefits from grant from the Association Nationale de la Recherche et de la Technologie, with Pierre Fabre, Convention industrielle de formation par la recherche number 2020/1701