Flexible Splines Modeling of Absolute Risks in Survival Analysis with R and SAS

Michael G. Palazzolo[1], Jeong-Gun Park[1], Giorgio EM. Melloni[1], Andrea Discacciati[2], Sabina A. Murphy[1], Andrea Bellavia[1]

[1]TIMI Study Group, Division of Cardiovascular Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA (US);
[2]Department of Medical Biostatistics, Karolinska Institutet, Stockholm (Sweden)

INTRODUCTION

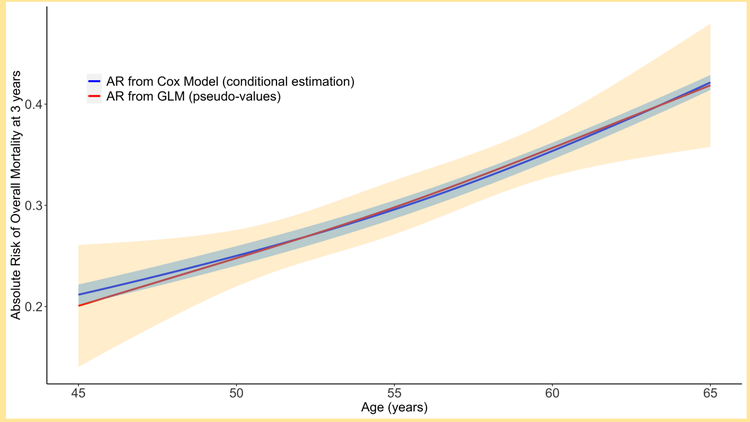
- Absolute Risks (AR, or event probabilities), represent a clinically relevant measure in survival analysis that should complement HRs in regular practice
- Splines transformations are recommended to incorporate potential non-linearities when evaluating continuous covariates in regression models
- We reviewed current approaches for estimating time-specific ARs from statistical models, and extended SAS and R material to flexibly account for non-linearities

STATISTICAL MODELING OF ABSOLUTE RISKS

- Statistical models for the hazard (e.g.: Cox, parametric survival) followed by absolute risk derivation through the survival-cumulative hazard relationship. Different approaches to handle covariates adjustment:
  - Inverse Probability Weights
  - Regression Standardization[1]
  - Conditional estimation
- Direct modeling of the absolute risk:
  - Generalized Linear Models using pseudo-values[2]
  - Risk Regression[3]

ILLUSTRATIVE EXAMPLE

- Simulated data on 10,000 individuals with 3 years of follow-up. Data based on Weibull distribution; age has a log-linear effect while BMI has a quadratic effect on overall mortality
- The figures show an example of splines modeled using a flexible display of time-specific AR for both a linear (panel A) and non-linear (panel B) example. Scan QR code for simulation details
- SAS macro and R code for flexible estimation of ARs available online (scan QR code)



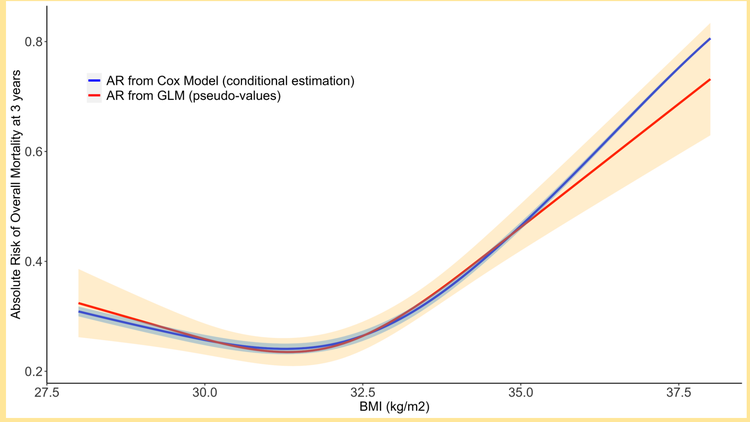


Figure. Absolute Risk of Overall Mortality at 3 years over levels of age (upper panel) and BMI (lower panel), modeled with restricted cubic splines in a Cox model (blue lines) and GLM model with pseudo-values (red line), in a simulated population.

Table: R and SAS functions to estimate Absolute Risks from statistical models

Risk prediction after regression modeling of the hazard	**R:** • `predict` functions from `survival` and `rms` packages • `predictSurvProb` from `pec` package • `stdReg`[1]
	SAS: • PROC PHREG (BASELINE) • %ANALY_PHREG_RCS*
Direct modeling of absolute risk	**R:** • `eventglm` (GLM with PV)[2] • `riskRegression`[3]
	SAS: • Original macro for pseudo values[4] • Extension: %ANALY_PSEUDO_RCS*

* SAS macros developed by the Authors and downloadable using the QR code link

CONCLUSIONS

- Several R packages are available to estimate AR with different modeling techniques
- We extended some of the available software to include splines modeling and flexible display of AR, and developed a new set of SAS macros
- Future work will include incorporating interactions with flexible transformation and their estimation on the risk scale
- Tools are available to present results in terms of both hazard and risk after multivariable adjustment in survival analysis

Contact: mgpalazzolo@bwh.harvard.edu
timi.org/biostatistics

References:
1 Sjölander A. Estimation of causal effect measures with the R-package *stdReg*. EJE. 2018.
2 Gerds TA et al. Absolute risk regression for competing risks: interpretation, link functions, and prediction. Stat in medicine. 2012.
3 Sachs MC, Gabriel EE. Event history regression with pseudo-observations. Journal of Statistical Software. 2022
4 Klein JP, et al.. SAS and R functions to compute pseudo-values for censored data regression. Computer methods and programs in biomedicine. 2008