

STRATOS Survival Analysis Subgroup

Terry Therneau

Mayo Clinic

August 2015

Current survival analysis group

- ▶ Michal Abrahamowicz
- ▶ Per Kragh Andersen
- ▶ Richard Cook
- ▶ Pierre Joly
- ▶ Torben Martinussen
- ▶ Maja Pohar-Perme
- ▶ Terry Therneau

There is obvious overlap with the initial STRATOS topic groups:

- ▶ TG1: missing data
- ▶ TG2: selection of variables and functional form
- ▶ TG3: descriptive and initial data analysis
- ▶ TG4: measurement error and misclassification
- ▶ TG5: study design
- ▶ TG6: evaluating diagnostic tests and prediction models
- ▶ TG7: causal models

- ▶ TG1: missing data
- ▶ TG2: selection of variables and functional form
- ▶ TG3: descriptive and initial data analysis
- ▶ TG4: measurement error and misclassification
- ▶ TG5: study design
- ▶ TG6: evaluating diagnostic tests and prediction models
- ▶ TG7: causal models

- ▶ TG1: missing data
- ▶ TG2: selection of variables and functional form
 - ▶ time-dependence and non-linear effects
- ▶ TG3: descriptive and initial data analysis
- ▶ TG4: measurement error and misclassification
- ▶ TG5: study design
- ▶ TG6: evaluating diagnostic tests and prediction models
- ▶ TG7: causal models

- ▶ TG1: missing data
- ▶ TG2: selection of variables and functional form
- ▶ TG3: **descriptive and initial data analysis**
 - ▶ survival curves
 - ▶ event rates and person-years
- ▶ TG4: measurement error and misclassification
- ▶ TG5: **study design**
 - ▶ sample size for survival studies
 - ▶ special designs
- ▶ TG6: evaluating diagnostic tests and prediction models
- ▶ TG7: causal models

- ▶ TG1: missing data
- ▶ TG2: selection of variables and functional form
- ▶ TG3: descriptive and initial data analysis
- ▶ TG4: measurement error and misclassification
- ▶ TG5: study design
- ▶ TG6: evaluating diagnostic tests and prediction models
- ▶ TG7: **causal models**
 - ▶ ipw and marginal-structural models

Specific topics

- ▶ Time dependent covariates: usage and pitfalls
- ▶ Time-dependent effects and flexible modeling
- ▶ Multiple events
 - ▶ competing risks and multi-state models
 - ▶ recurrent events
 - ▶ joint models of survival and longitudinal markers
- ▶ Particular models
 - ▶ Cox proportional hazards
 - ▶ additive
 - ▶ accelerated failure time, parametric and non-parametric
 - ▶ restricted mean life
- ▶ Penalized models and random effects
- ▶ Relative survival
- ▶ Interval censoring
- ▶ Causal/cumulative effects
- ▶ Validation

- ▶ Time dependent covariates

- ▶ Time dependent covariates
 - ▶ survival by treatment response, immortal time bias
 - ▶ reverse causality, end of life markers
 - ▶ use of midpoints for interval censored data
 - ▶ ...

- ▶ Time dependent covariates
 - ▶ survival by treatment response, immortal time bias
 - ▶ reverse causality, end of life markers
 - ▶ use of midpoints for interval censored data
 - ▶ ...
- ▶ Time dependent effects
 - ▶ $\beta(t)x$ NOT $\beta x(t)$

- ▶ Multiple events
 - ▶ Medical need

- ▶ Multiple events
 - ▶ Medical need
 - ▶ competing risks and multi-state
 - ▶ repeated events
 - ▶ joint modeling

- ▶ Particular models
 - ▶ Cox model

- ▶ Particular models
 - ▶ Cox model
 - ▶ Additive hazard

▶ Particular models

▶ Cox model

▶ Additive hazard

▶ $\lambda_1(t)e^{X\beta} + \lambda_2(t)e^{X\gamma} + \dots$

▶ Particular models

- ▶ Cox model
- ▶ Additive hazard
 - ▶ $\lambda_1(t)e^{X\beta} + \lambda_2(t)e^{X\gamma} + \dots$
- ▶ Accelerated failure time
- ▶ Residual life

▶ Penalized models and random effects

▶ $\lambda_0(t)e^{X\beta+Zb}$

$Db \sim G(0, s)$

▶ Classic mixed effects: $D = I$, $G =$ Gaussian, general s

▶ Penalized models and random effects

▶ $\lambda_0(t)e^{X\beta+Zb}$

$Db \sim G(0, s)$

▶ Classic mixed effects: $D = I$, $G =$ Gaussian, general s

▶ Lasso, Elastic net, logistic

▶ Penalized models and random effects

- ▶ $\lambda_0(t)e^{X\beta+Zb}$
 $Db \sim G(0, s)$
- ▶ Classic mixed effects: $D = I$, $G =$ Gaussian, general s
- ▶ Lasso, Elastic net, logistic
- ▶ kill use of the word 'frailty'

- ▶ Penalized models and random effects
 - ▶ $\lambda_0(t)e^{X\beta+Zb}$
 $Db \sim G(0, s)$
 - ▶ Classic mixed effects: $D = I$, $G =$ Gaussian, general s
 - ▶ Lasso, Elastic net, logistic
 - ▶ kill use of the word 'frailty'
- ▶ Relative survival
- ▶ Interval censoring
- ▶ Causal modeling

- ▶ Software validation
- ▶ Many of these are complex methods

- ▶ Software validation
- ▶ Many of these are complex methods
- ▶ I no longer trust unvalidated survival software

- ▶ Software validation
- ▶ Many of these are complex methods
- ▶ I no longer trust unvalidated survival software
- ▶ R survival
 - ▶ 29K code, 12K test, (7 + 7 + 4) documentation
 - ▶ any untested code has a 1 in 5 chance of containing an error

- ▶ Software validation
- ▶ Many of these are complex methods
- ▶ I no longer trust unvalidated survival software
- ▶ R survival
 - ▶ 29K code, 12K test, (7 + 7 + 4) documentation
 - ▶ any untested code has a 1 in 5 chance of containing an error
- ▶ What is a sufficient test suite?

Current survival analysis group

- ▶ Michal Abrahamowicz
- ▶ Per Kragh Andersen
- ▶ Richard Cook
- ▶ Pierre Joly
- ▶ Torben Martinussen
- ▶ Maja Pohar-Perme
- ▶ Terry Therneau