Weighted Estimation in Cox Regression Revisited

Michael Schemper, Samo Wakounig and Georg Heinze

Section of Clinical Biometrics
Core Unit of Medical Statistics and Informatics
Medical University of Vienna, Austria

Project sponsored by the Austrian Research Fund

Contents

- Average hazard ratios in a population
- Weighted estimation in Cox regression
- Time - dependent effects and residuals for weighted estimation
- Comparison of standard and weighted estimation in Monte Carlo study
- Comparative statistical analyses of PBC study

Motivation

- Analysis by the proportional hazards (Cox) model if hazards are not proportional leads to
  - biased estimates of the (average) hazard ratio (AHR)
  - possible loss of power of tests of the AHR
- Methods by additional parameters available
- Explicit estimation of the AHR avoids
  - the disadvantages of the standard Cox model
  - and the need for additional parameters

⟶ good choice with small samples and/or many covariates

AHR in Population (1)

- Possible definitions of a *HR* on continuous time:

$$\exp\left[\int \log\left(h_0(t)/h_1(t)\right) f(t)\ dt\right]$$ *'intuitive'*

versus

'pragmatic' (used in Cox model and by Mantel-Haenszel est.)

$$\frac{\int \left(h_0(t)/h(t)\right) f(t)\ dt}{\int \left(h_1(t)/h(t)\right) f(t)\ dt}$$

where $h_0(t)$ and $h_1(t)$ denote the hazards of groups G_0 and G_1, respectively $h(t) = h_0(t) + h_1(t)$
$f(t)$ is uniform;

AHR in Population (2)

- Definition of a *HR* in Cox‘s philosophy:

$$\frac{\int \left(h_0(t)/h(t)\right)\ f(t)\ dt}{\int \left(h_1(t)/h(t)\right)\ f(t)\ dt}$$

where $f(t)$ denotes the frequency (density) of events (i.e., tables in the model) at *t.*

$$f(t) = -\left(\sqrt{S_0(t)\ S_1(t)}\right)'$$

AHR in Population (3)

- Definition of the *pragmatic* ***AHR*** within Cox‘s philosophy:

$$\frac{\int \left(h_0(t) / h(t) \right) f(t)\, S(t)\, dt}{\int \left(h_1(t) / h(t) \right) f(t)\, S(t)\, dt}$$

where $S(t)$ denotes the overall survival function, symbolizing the number of individuals affected by the hazard ratio at time ***t***.

This ***AHR*** *is estimated by the weighted Cox estimation we propose.*

AHR in Population (4)

Is the *intuitive* $AHR_{int} = \exp\left\{\frac{\int \log\left(h_0(t)/h_1(t)\right) f(t)\,S(t)\,dt}{\int f(t)\,S(t)\,dt}\right\}$

similar

to the *pragmatic* $AHR_{pr} = \frac{\int \left(h_0(t)/h(t)\right) f(t)\,S(t)\,dt}{\int \left(h_1(t)/h(t)\right) f(t)\,S(t)\,dt}$?

Response:

- For proportional hazards both definitions are identical
- For non-proportional hazards they are almost identical for realistic scenarios. ⟶ Intuitive definition can be used!

AHR in Population (5)

Note that the $\boldsymbol{AHR}_{pr} = \dfrac{\int \left(h_0(t) / h(t) \right) \, f(t) \, S(t) \, dt}{\int \left(h_1(t) / h(t) \right) \, f(t) \, S(t) \, dt}$

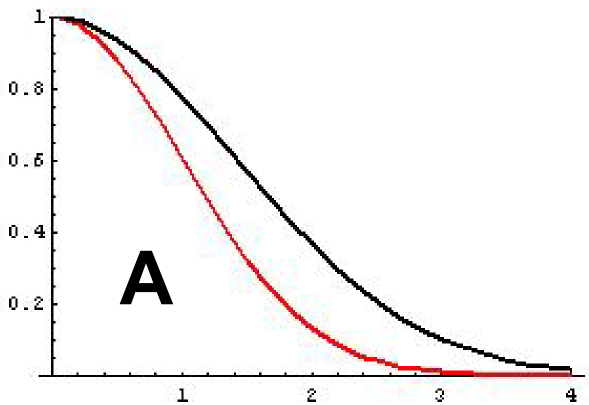
with $S(t) = \sqrt{S_0(t) \; S_1(t)}$

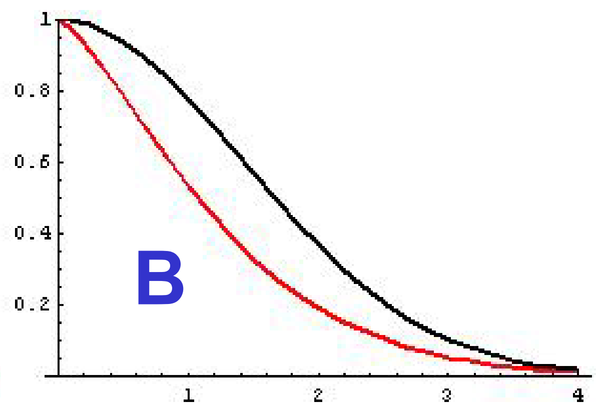
is mathematically identical to the hazard ratio definition

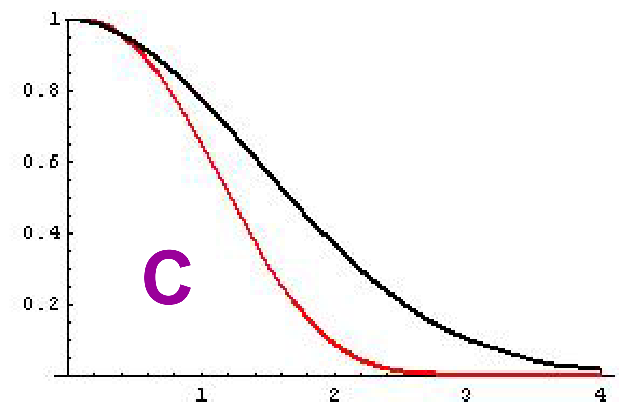
$$ORES = P(T_0 < T_1) / (1 - P(T_0 < T_1))$$

„**Odds Ratio of Effect Size**“, which also does not require proportional hazards and gives equal weight to all individuals (by pairwise comparisons of all times T_0 and T_1).

Empirical Investigations (Weibull-Populations)







Hazard ratios in populations	A	B	C
ORES	2.0	2.0	2.0
w. Cox *intuitive*	2.0	2.1	2.0
w. Cox *pragmatic*	2.0	2.0	2.0
Cox	2.0	1.5	2.4

Weighted estimation in Cox regression (1)

Sample of n individuals:

- m uncensored survival times t_j among n possibly censored survival times; survival status; $\longrightarrow$ risk sets R_j
- k covariate values for each individual

Then for each of the k covariates, say the r th, the following estimating equation is defined:

$$\sum_{j=1}^{m} w_r(t_j) \left[x_{jr} - \frac{\sum_{l \in R_j} x_{lr} \exp(x_l \hat{\beta})}{\sum_{l \in R_j} \exp(x_l \hat{\beta})} \right] = 0$$

weight at t_j

observed – expected covariate value

Weighted estimation in Cox regression (2)

$$\sum_{j=1}^{m} w_r(t_j)\left[x_{jr} - \frac{\sum_{l \in R_j} x_{lr} \exp(x_l \hat{\beta})}{\sum_{l \in R_j} \exp(x_l \hat{\beta})} \right] = 0$$

- $\hat{\beta}$ obtained as solutions to this equation.
- If $w_r(t_j) = 1$ $\longrightarrow$ standard Cox model.
- Other choices for $w_r(t_j)$ possible:
  - $w_r(t_j) = R_j$ $\longrightarrow$ generalisation of Breslow (1974) - test
  - $w_r(t_j) = S_j$ $\longrightarrow$ generalisation of Prentice (1978) - test

  to multiple regression models.

Weighted estimation in Cox regression (3)

- The choices of $w_r(t_j)$ reflect the relative importance attached to hazard ratios at different times; $w_r(t_j) = R_j$ and $w_r(t_j) = S_j$ weight by the number of individuals actually or likely affected by the log hazard ratio β at t_j.

- This weighting gives equal weight to all individuals in the case of no censoring and the resulting $\hat{\beta}$ does not rely on the assumption of proportional hazards.

Weighted estimation in Cox regression (4)

- Normalization of weights $w_r(t_j)$ required for some inference procedures:

$$w_r^n(t_j) = w_r(t_j)\left[\sum_{j=1}^{m} R_j \Big/ \sum_{j=1}^{m} R_j\, w_r(t_j)\right]$$

Note: the normalization of the weights leaves the relative contribution to the partial likelihood at t_j unchanged.

Tests and Confidence Intervals

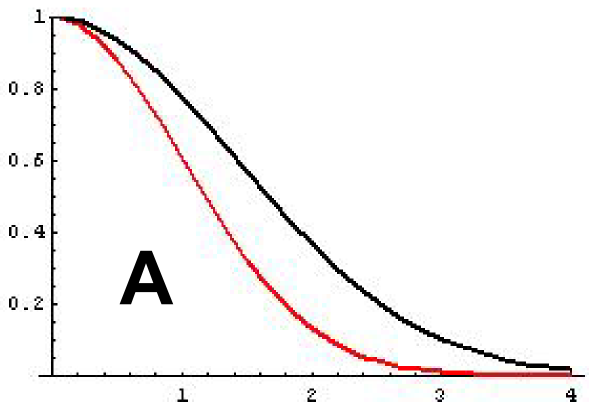
- Wald* (via covariance matrix from information matrix)

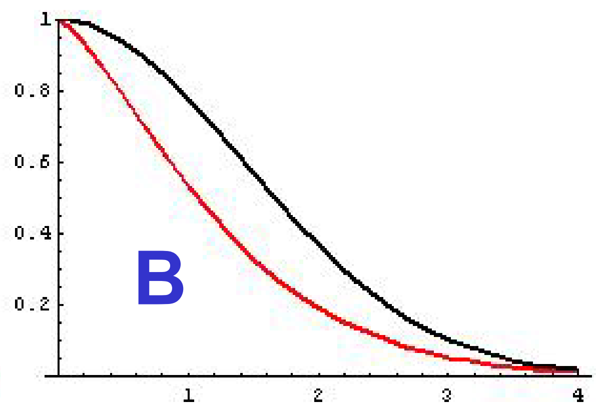
$$I_{rs}(\hat{\beta}) = \frac{-\partial^2 \log L}{\partial \beta_r \, \partial \beta_s} = \sum_{j=1}^{m} w_r^n(t_j)\, w_s^n(t_j) \{ \ldots \}$$

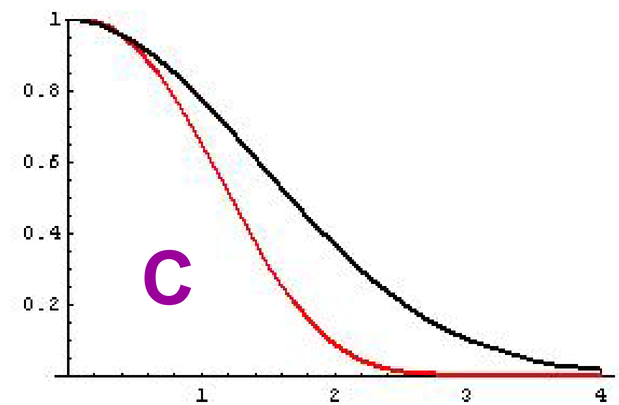
- Wald (via robust ‘sandwich‘ estimate of the covariance matrix)
- Score* (confidence intervals numerically more intensive)
- Likelihood ratio (only for equal weights for all covariates)

* implemented in software

Empirical Investigations (Weibull-Populations)

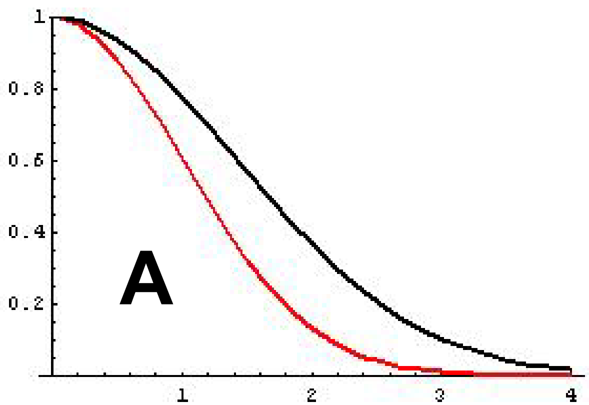


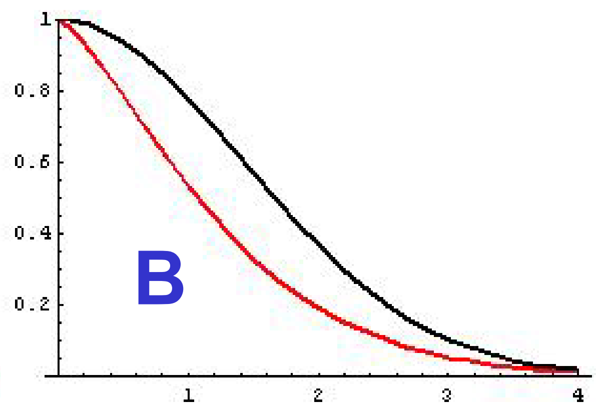


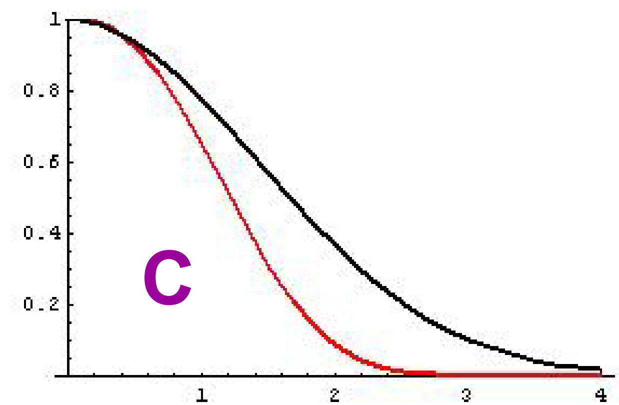


Hazard ratios in populations	A	B	C
ORES	2.0	2.0	2.0
w. Cox	2.0	2.0	2.0
Cox	2.0	1.5	2.6

Empirical Investigations (Study of Bias)





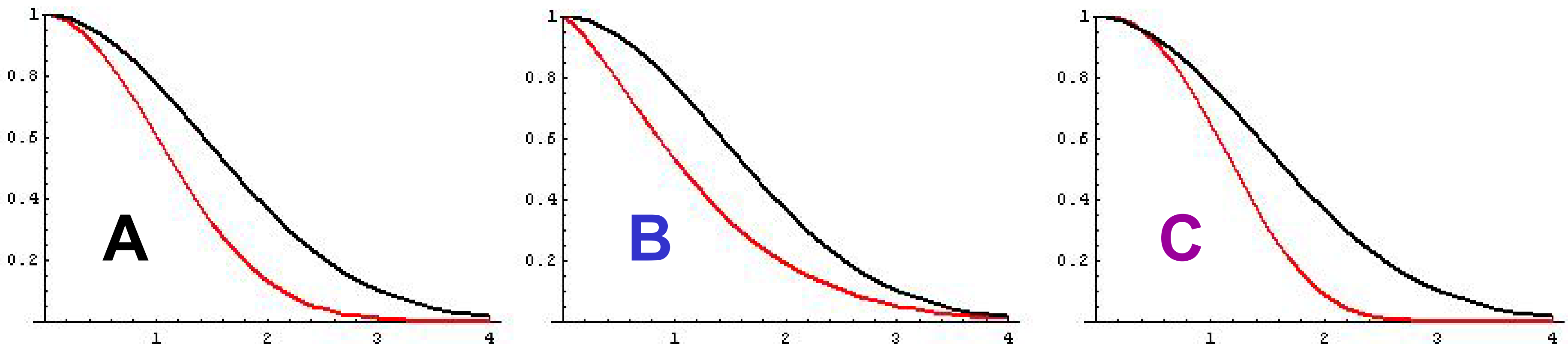


Bias of $\hat{\beta}$ $\beta = \log 2$	A	B	C
Cox	0.02	-0.21	0.15
w. Cox	0.01	0.00	0.03

Simulated samples: 10 000

$n_1 = n_2 = 40$

Empirical Investigations (Study of Precision)

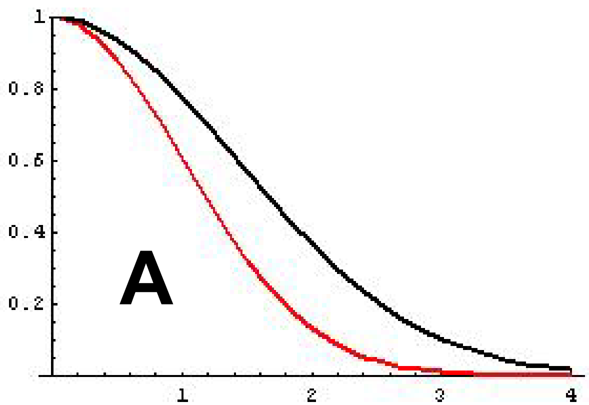


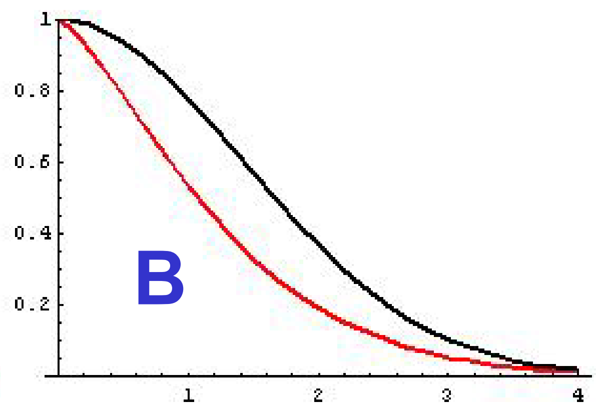
$med\left(std(\hat{\beta})\right)$	A	B	C
Cox	0.24	0.23	0.25
w. Cox	0.27	0.26	0.27

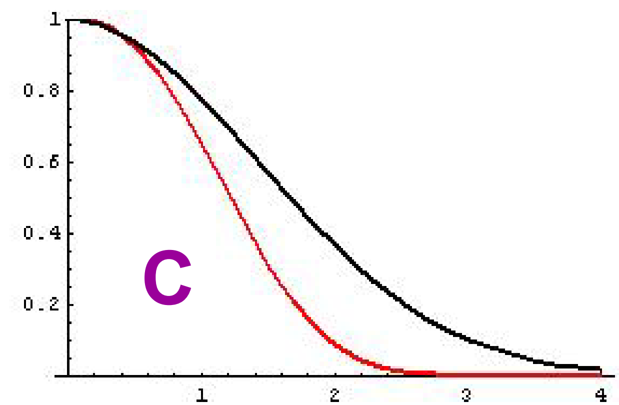
Simulated samples: 10 000

$n_1 = n_2 = 40$

Empirical Investigations (Study of Power)







Power of test $\beta = 0$	A	B	C
Cox	0.85	0.53	0.94
w. Cox	0.75	0.75	0.76

Simulated samples: 10 000

$n_1 = n_2 = 40$

Score, Wald and robust confidence intervals

- Simulation study:
  to compare coverage of confidence intervals by
  - Score approach
  - Wald (standard covariance matrix)
  - Wald (robust covariance matrix)
- Survival times follow exponential distribution
- Two balanced binary covariates x_1 and x_2
  - $\beta_1 = 0$, $\beta_2 = -1$
  - 10 000 simulated samples, $n_1 = n_2 = 40$

Coverage by confidence intervals (nominal 97.5%)

		Left-sided			Right-sided		
	Weights S_j applied to	**Score**	**Wald**	**Robust**	**Score**	**Wald**	**Robust**
$\beta_1 = 0$	None	96.8	97.0	96.6	97.0	97.2	97.2
	X_1	96.9	96.8		97.6	97.5	
	X_2	96.6	96.7		97.0	97.1	
	$X_1 + X_2$	96.8	96.7	96.6	97.4	97.4	97.3
$\beta_2 = -1$	None	97.5	97.6	97.1	95.2	96.9	96.3
	X_1	97.5	97.5		95.0	96.5	
	X_2	97.2	97.0		96.5	97.3	
	$X_1 + X_2$	97.3	97.0	97.1	96.4	97.1	96.3

Comparison of score, Wald and robust confidence intervals

Summary of results

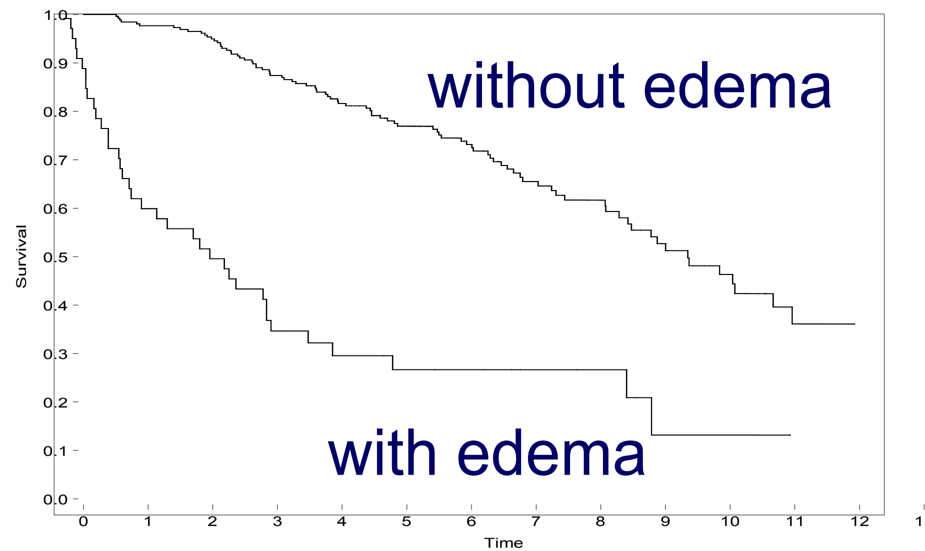
- With n=40 nominal coverage is slightly violated in almost all scenarios, but similarly for Cox and weighted Cox.
- With n=120 (not shown) there are almost no violations by either method.
- The quality of coverage is the same for all three inference methods.
- The quality of coverage is the same for Cox and weighted Cox.

Example: Primary Biliary Cirrhosis Trial

- Study of survival of n=312 patients of the Mayo Clinic, (60% censored survival times)
- Five prognostic factors used:
  - Age (in years)
  - Edema (no / yes)
  - log (Bilirubin)
  - log (Prothrombin time)
  - Albumin
- Edema now studied in more detail

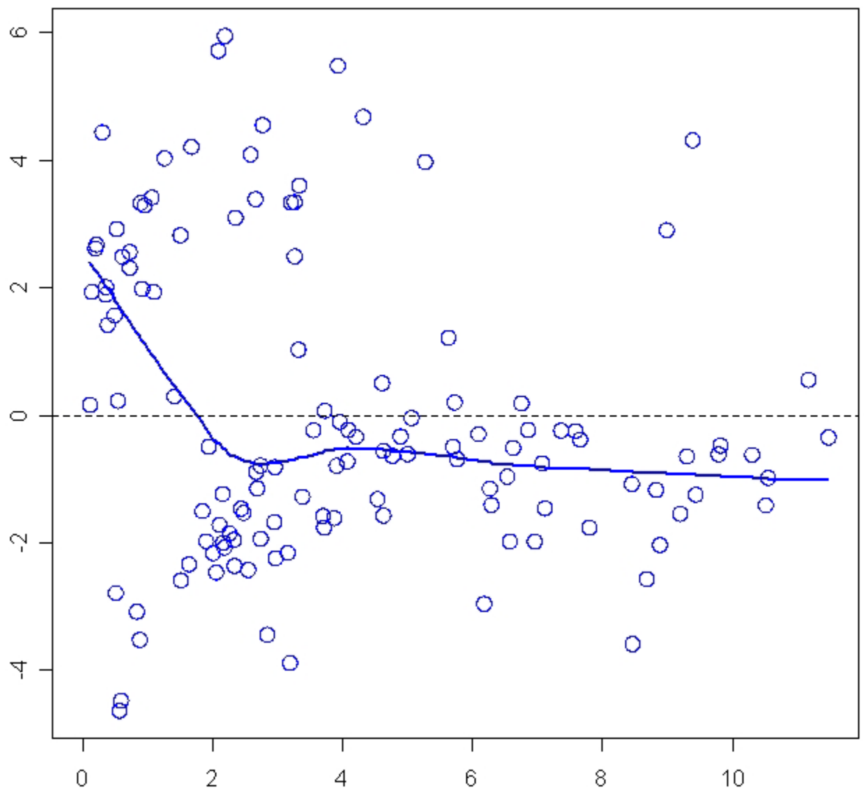
Primary Biliary Cirrhosis Trial

KM - survival functions and analysis by Cox regression

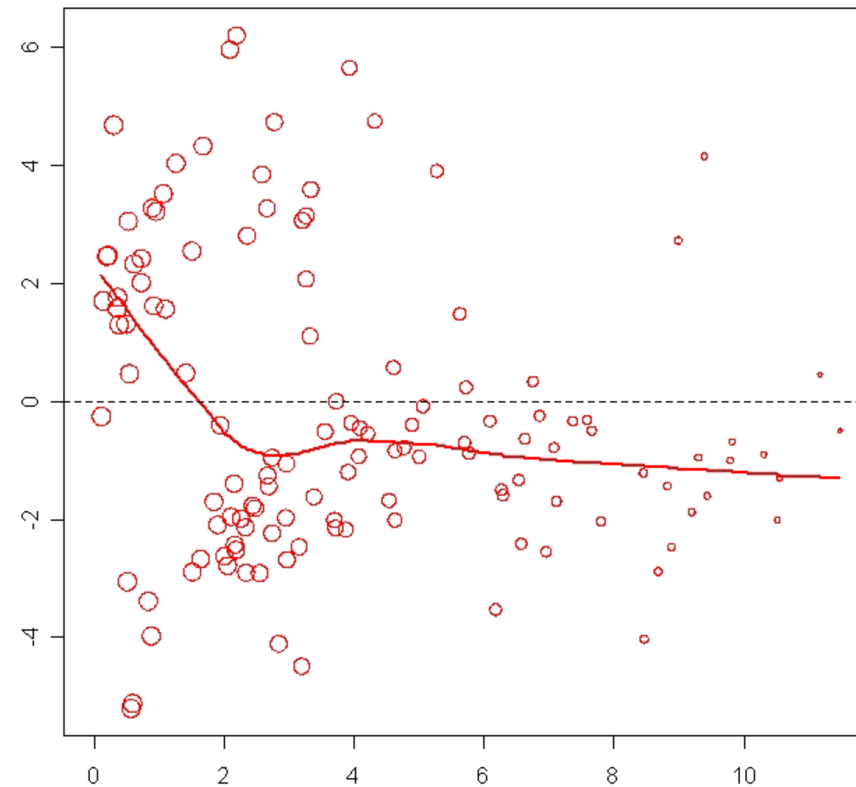


Hazard Ratios	Unweighted	Weighted (R_j)
Unadjusted	4.00 (p < 0.0001)	5.46 (p < 0.0001)
Adjusted	1.48 (p = 0.08)	1.81 (p = 0.009)

Schoenfeld Residual plots (weight R_j) for Edema

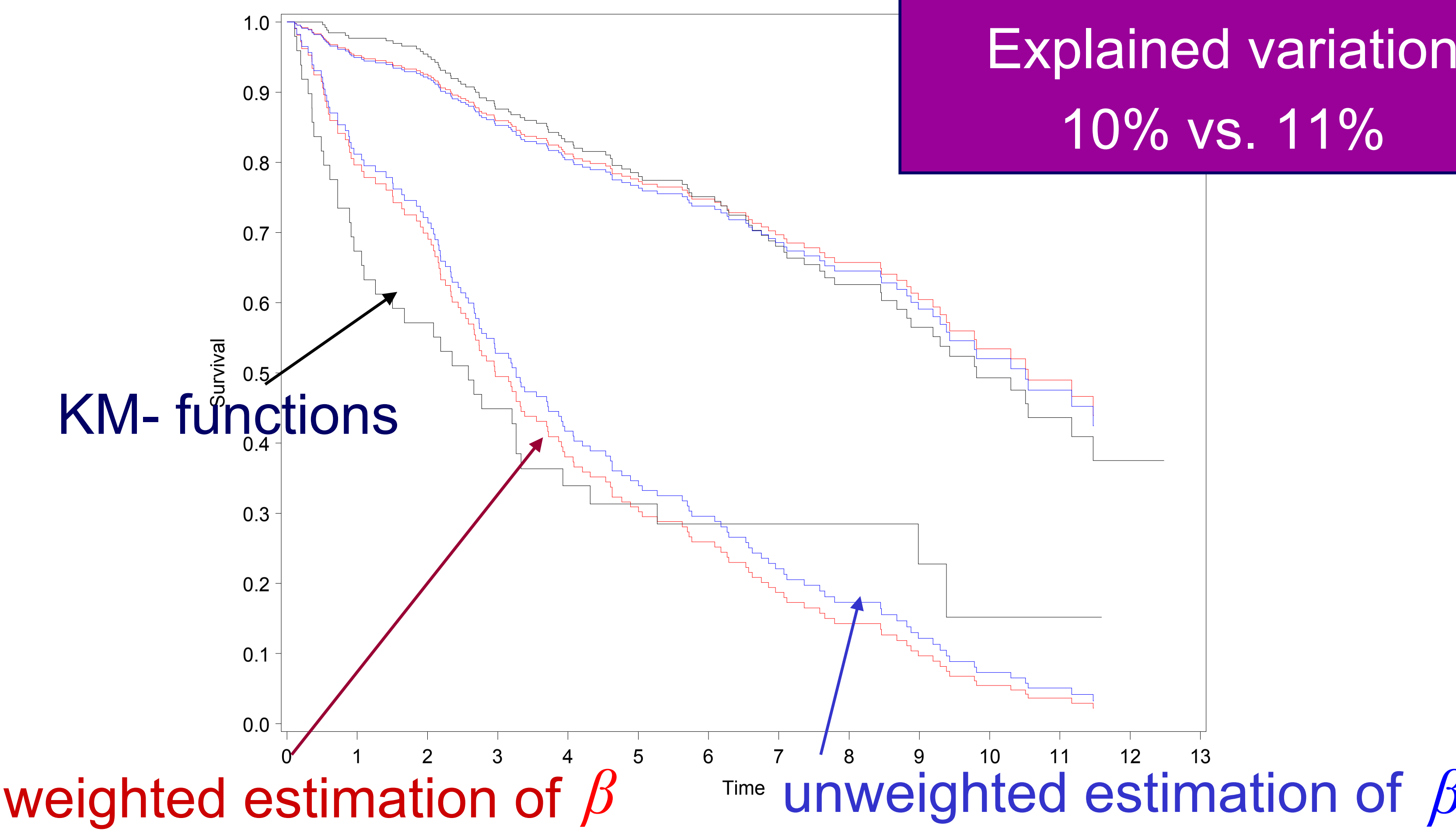


Standard plot for unweighted estimation



Modified plot with for weighted estimation (dots proportional to R_j)

Survival functions for Edema based on weighted versus unweighted estimation

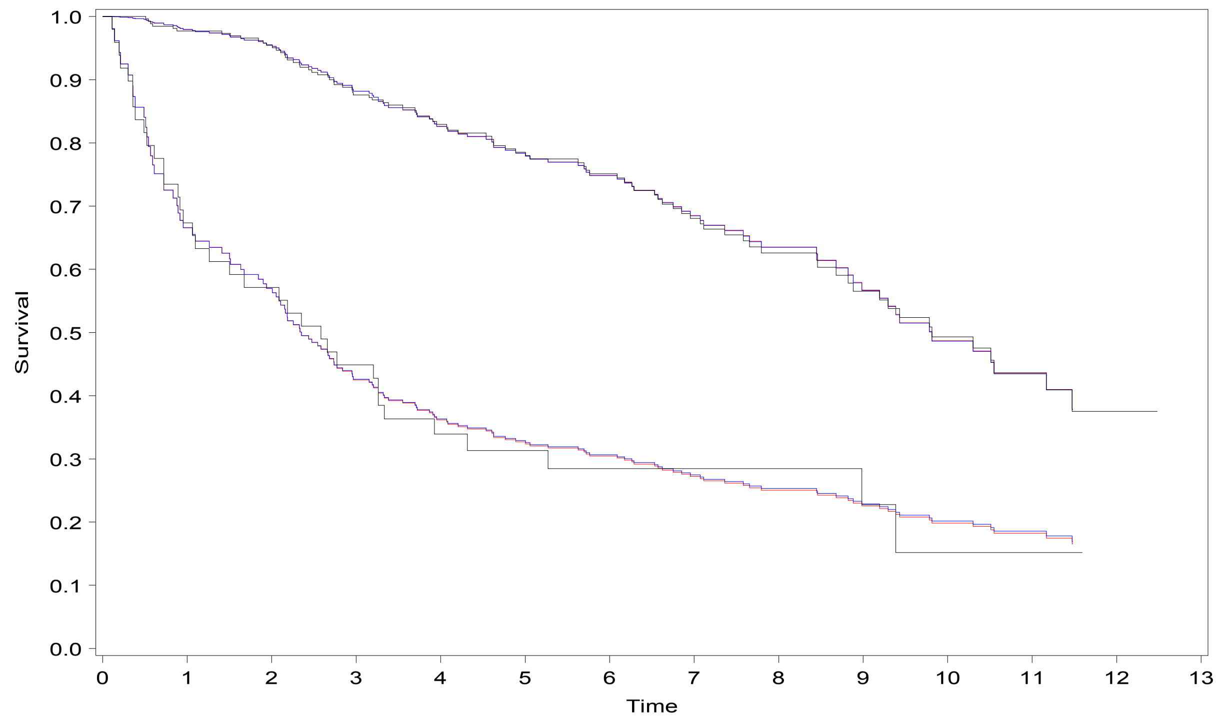


Explained variation

10% vs. 11%

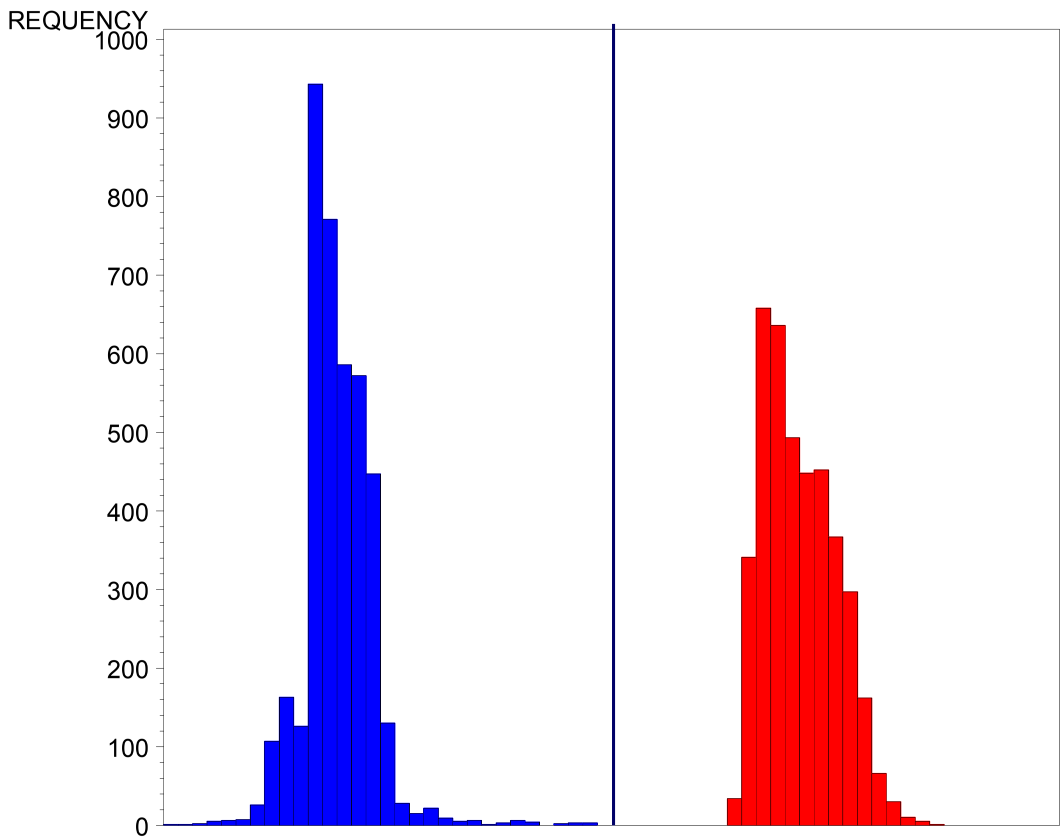
‘Time-dependent‘ Survival functions for Edema based on weighted versus unweighted estimation

KM- functions (black)



‘Time-dependent‘ survival functions (using $\beta x + \beta_t x \log t$) are virtually identical under unweighted and weighted estimation of β !

dfbetas for Edema
based on weighted versus unweighted estimation



unweighted estimation weighted estimation

Implementation of weighted estimation for Cox regression

We produced specialized software based on FORTRAN 90:

- SAS macro %WCM
- R package coxphw

<u>Some options:</u>

- Weights can be set to R_j, S_j or 1 for each covariate separately
- Inference (tests and CIs) based on Wald or score statistics
- Fixed and time-dependent effects
- Optional counting-process style input

Conclusions on the role of weighted estimation

- Disadvantages of weighted estimation:
  - $std\,(\hat{\beta})$ slightly larger
- Equal performance:
  - Time-dependent effects modelling
  - Explained variation
  - Coverage by confidence intervals
- Advantages of weighted estimation:
  - Hazard ratio estimates always interpretable
  - Decision - theoretic: $P(T_0 < T_1)$
  - Robustness (dfbetas)

Some references (chronologic order)

- Kalbfleisch, JD and Prentice, RL (1981). Estimation of the average hazard ratio. *Biometrika* **68**, 105-112.
- Lin, DY (1991). Goodness-of-fit analysis for the Cox regression model based on a class of parameter estimators. *JASA* **86**, 725-728.
- Schemper, M (1992). Cox analysis of survival data with non-proportional hazard functions. *The Statistician* **41**, 455-465.
- Sasieni, P (1993). Maximum weighted partial likelihood estimators for the Cox model. *JASA* **88**, 144-152.
- Minder, CE and Bednarski, T (1996). A robust method for proportional hazards regression. *Stat.Med.* **15**, 1033-1047.
- Valsecchi, MG, Silvestri D and Sasieni P (1996). Evaluation of long-term survival: use of diagnostics and robust estimators…*Stat.Med*. **15**, 2763-2780.
- Schemper, M, Wakounig S and Heinze G. *In preparation*.

My colleagues of the Weighted Cox Project

Georg Heinze

Samo Wakounig

Daniela Dunkler

Programs in SAS and R

http://www.muw.ac.at/msi/biometrie