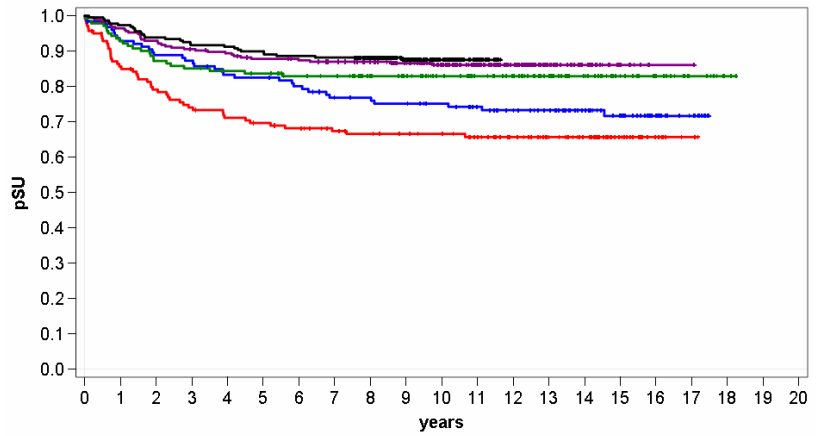
A new approach based on pseudo-values for assessing the effect of a binary time-dependent covariate on patient survival

U. Pötschger, H. Heinzl, M. G. Valsecchi, A. Attarbaschi, M. Minkov, M. Mittlböck

Outline

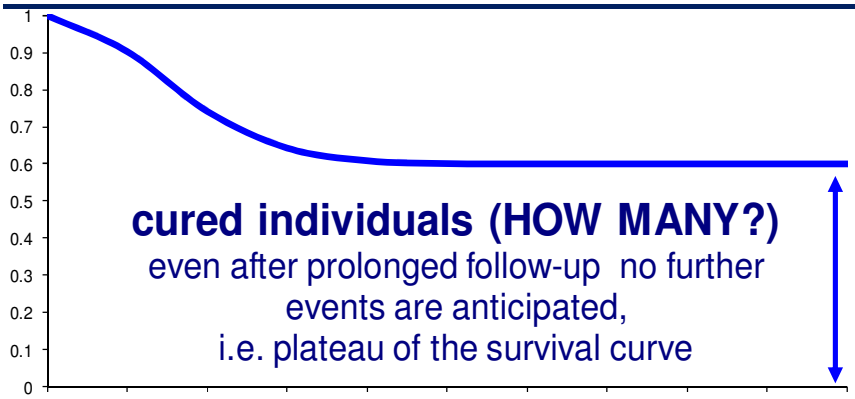
- **Paediatric oncology:**
  Survival endpoint; aim is to improve long-term outcome
- **Pseudo-value regression technique**
- **Methodological challenges when comparing SCT and chemo-therapy**
- **A novel statistical approach for investigating the impact of a binary time-dependent covariates on long-term outcome**

Childhood leukemia



ALL 81 ALL84 ALL86 ALL90 ALL95

Typical survival function



cured individuals (HOW MANY?)

even after prolonged follow-up no further events are anticipated, i.e. plateau of the survival curve

time to event (WHEN?)

Cox regression

addresses

average hazard ratios

relies on proportional hazards and
does not distinguish between
1) failure time
2) long-term failure rates

Proper methods

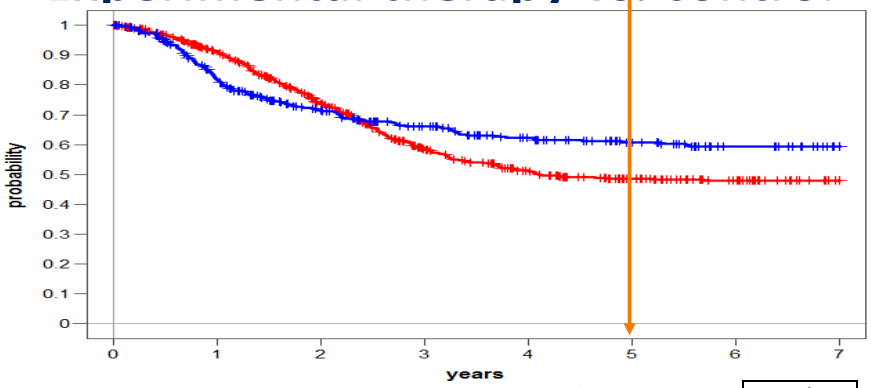
should address

primary aim
long-term outcome

Primary Aim

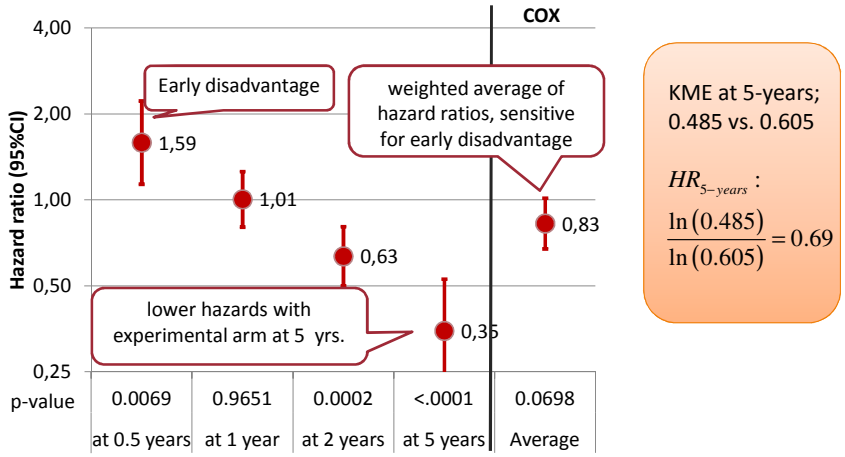
Example 1:

Experimental therapy vs. control



	pts	events	5-years p	p-value (log-rank)
Control (less intense)	590	256	0.49±0.02	0.069
Experimental (more intense)	410	143	0.61±0.03	

Example 1: Cox-regression



KME at 5-years; 0.485 vs. 0.605

$$HR_{5-years}: \frac{\ln(0.485)}{\ln(0.605)} = 0.69$$

Pseudo-value regression

- investigates survival rates at pre-specified time-points, e.g. **5-years survival rates**
- **ignores early differences**
- **no proportional hazard assumption**

Pseudo-values in survival analysis

$$\hat{V}_i(t) = n\hat{S}(t) - (n-1)\hat{S}^{-i}(t)$$

Notation:

t	Pre-specified time-point since time 0
n	Number of observations
$\hat{S}(t)$	Kaplan-Meier estimate for t-years survival (KME)
$\hat{S}^{-i}(t)$	KME, i-th patient left out

Example 1: Pseudo-Values

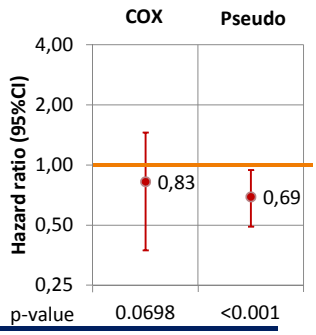
	Control			Experimental		
	Pseudo-value	Event-time	Event	Pseudo-value	Event-time	Event
	1.12	>5	0.00	**1.09**	4.17	0.00
	0.54	0.11	0.00	**-0.38**	4.06	1.00
	1.12	>5	0.00	**1.12**	>5	0.00
	-0.33	3.73	1.00	**0.70**	1.61	0.00
	-0.25	3.04	1.00	**1.12**	>5	0.00
	-0.24	2.97	1.00	**1.12**	>5	0.00
	1.12	>5	0.00	**-0.07**	1.37	1.00
	0.92	2.95	0.00	**-0.34**	3.79	1.00
	-0.07	1.31	1.00	**1.12**	>5	0.00
	…	…	…	**…**	…	…
n	**590**			**410**		
mean	**0.485**			**0.605**		
min	**-0.533**			**-0.607**		
max	**1.117**			**1.117**		
KME	**0.485**			**0.605**		

Generalised linear Model

- **Response**:
  n pseudovalues
- **Response probability distribution**:
  Normal distribution
- **Independent variables**
  p (>=1) factors + intercept
- **Link function g():**
  e.g. log(-log()), logit(), ...
- **Variance estimation:**
  empirical sandwich estimator

Example 1: Pseudo-value regression

Parameter			p-value		95% CI	
Therapy	→	HR	<0.001	0.694	0.561	0.860
Intercept (Control)		5-years p		0.485	0.438	0.530
Experimental		5-years p		0.605	0.551	0.654



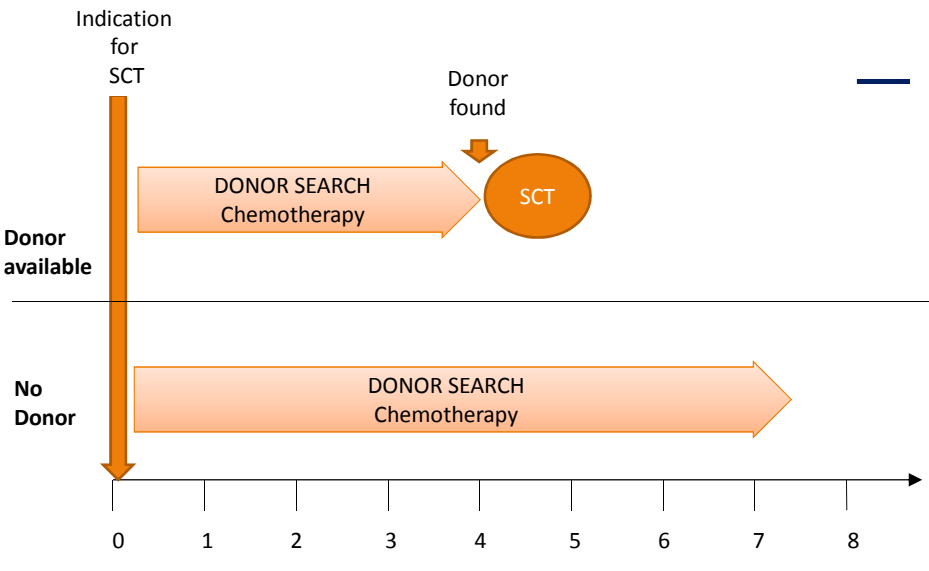
Motivating Example: PH+ ALL

Philadelphia chromosome positive acute lymphoblastic leukemia in children [Arico, JCO 2010]

Therapeutic options:

- Conventional chemotherapy
- Stem-cell transplantation when a matched donor can be found

Introduction Pseudo-values SCT-Chemo Pseudo-TDC Conclusion

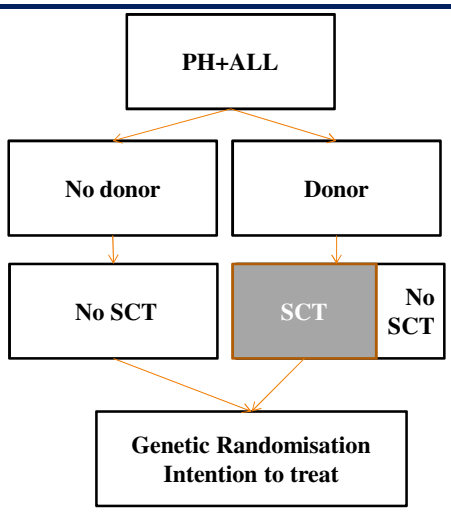


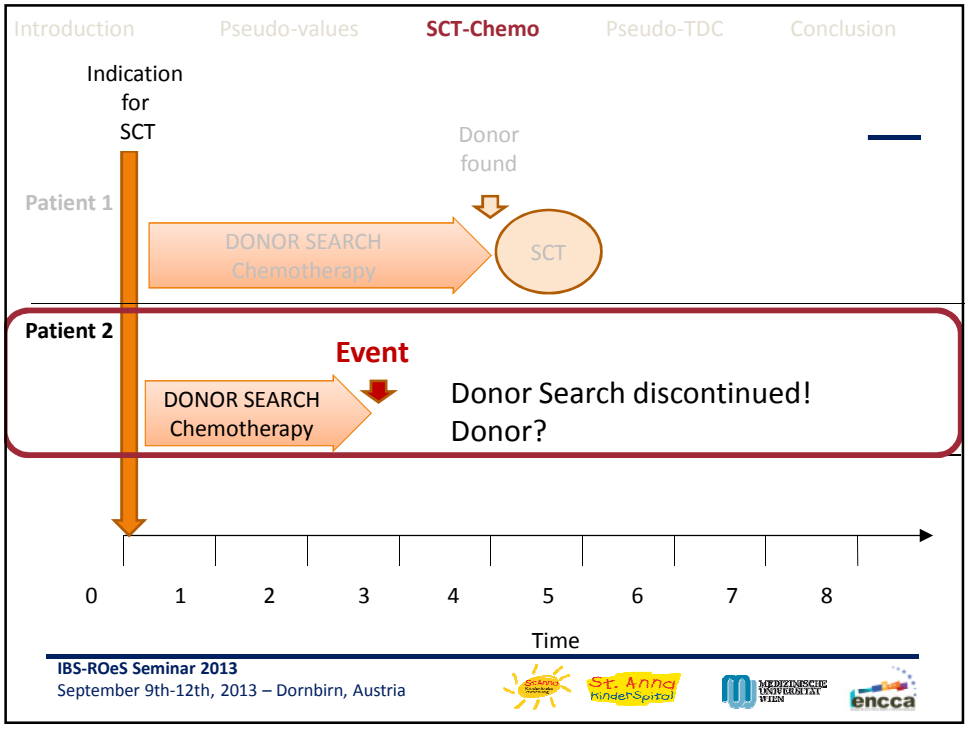
IBS-ROeS Seminar 2013
September 9th-12th, 2013 – Dornbirn, Austria

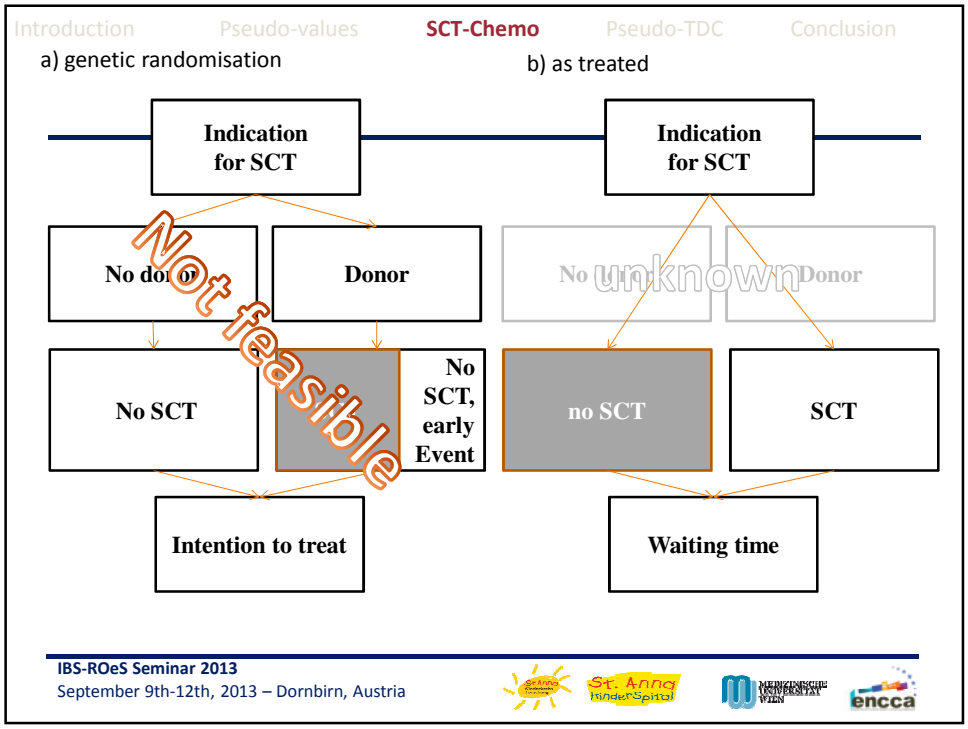
Motivating Example: PH+ ALL

- **Aim of the study:**
  Compare survival of stem-cell transplantation (SCT) versus chemotherapy
- **Question of the treating physician:**
  Should I perform an SCT, when a donor becomes available?

Genetic Randomisation




Introduction
Pseudo-values
SCT-Chemo
Pseudo-TDC
Conclusion
Indication for SCT
Donor found
Patient 1
DONOR SEARCH Chemotherapy
SCT
Patient 2
Event
DONOR SEARCH Chemotherapy
Donor Search discontinued!
Donor?
0
1
2
3
4
5
6
7
8
Time
IBS-ROeS Seminar 2013
September 9th-12th, 2013 – Dornbirn, Austria

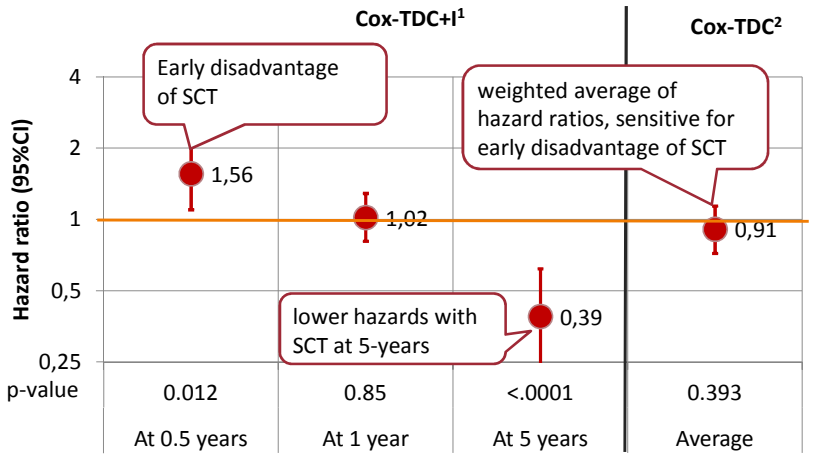

Introduction
Pseudo-values
SCT-Chemo
Pseudo-TDC
Conclusion
a) genetic randomisation
b) as treated
Indication for SCT
Indication for SCT
No donor
Donor
No unknown Donor
Not feasible
No SCT
No SCT, early Event
no SCT
SCT
Intention to treat
Waiting time
IBS-ROeS Seminar 2013
September 9th-12th, 2013 – Dornbirn, Austria

Conventionally used approaches

- Cox-regression with time-dependent covariates
- With non-proportional hazards:
  +/- interaction with time
- Landmark analysis
  mainly for estimation of survival rates
  - Classical approach [Anderson 1983]
  - Extended Kaplan-Meier estimates

Introduction Pseudo-values SCT-Chemo Pseudo-TDC Conclusion

Cox-regression with time-dependent covariates



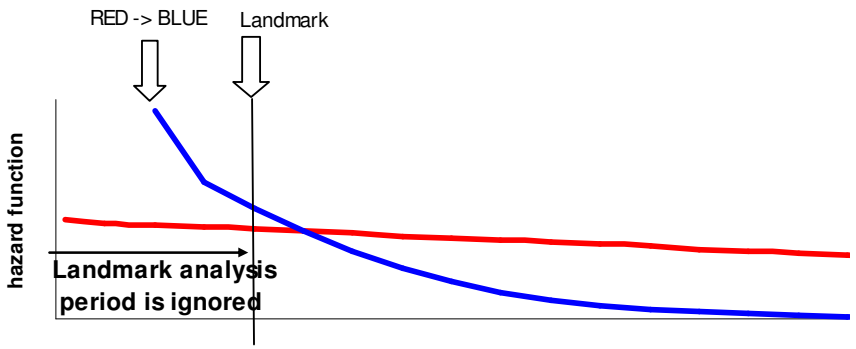
IBS-ROeS Seminar 2013
September 9th-12th, 2013 – Dornbirn, Austria

Conventionally used approaches

- Cox-regression with time-dependent covariates +/- interaction with time
- Landmark analysis
  mainly for estimation of survival rates
  - Classical approach [Anderson 1983]
  - Extended Kaplan-Meier estimates

Introduction Pseudo-values SCT-Chemo Pseudo-TDC Conclusion

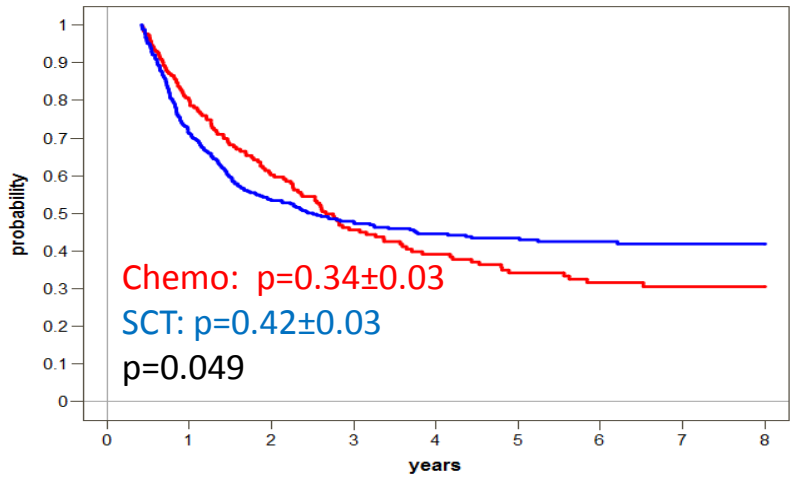
Estimation of survival-curves with TDC



- an **arbitrary landmark time** (after time 0) is needed
- survival curves start at landmark time
- several different methods are available
- results depend on the chosen landmark time

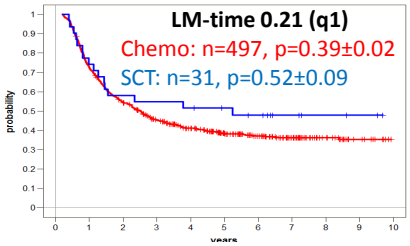
Landmark Analysis

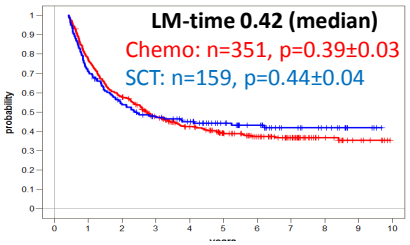
KM-extented (landmark time 0.42 median)



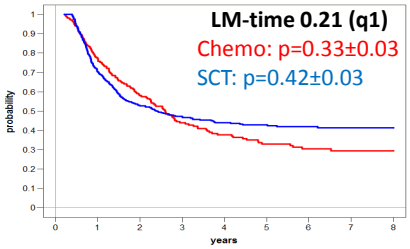
Landmark Analysis

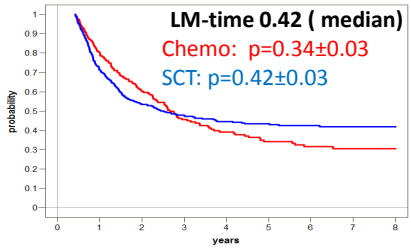
Anderson





KM-Ext



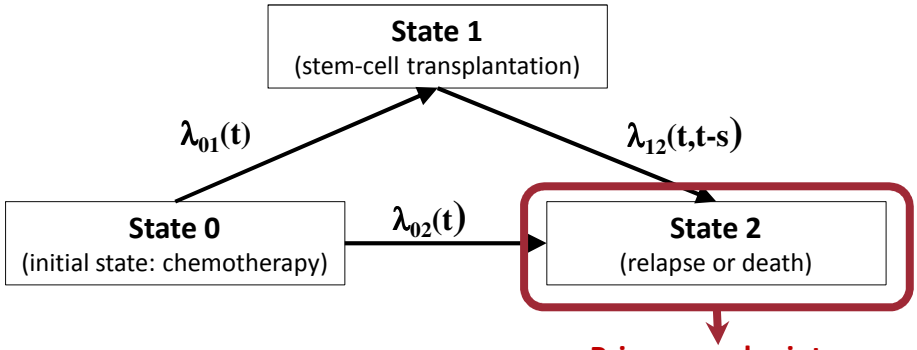


Time-dependent covariates and non-proportional hazards

Problem	Cox	Pseudo	LM	novel
PH-assumption not needed	×	✓	?	Needed
addresses long-term outcome	×	✓	✓	
adjustment for waiting-time bias	✓	×	✓	
no arbitrary specification landmark-time needed	✓	×	×	
inclusion of covariates	✓	✓	×	
parameter estimates clear interpretation	✓	✓	×	

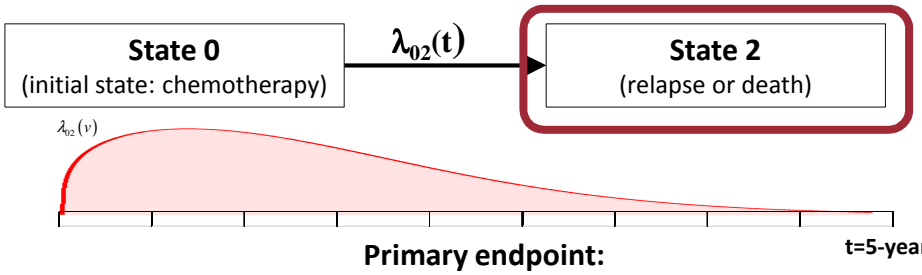
best methodological approach is unclear?
Novel approaches are needed !

3-state stochastic process



Primary endpoint:
Probability to reach state 2
in a given time-intervall

2-state stochastic process



Primary endpoint:
Probability to reach state 2
in a given time-intervall

$$S_0(t) = \exp\left(-\int_0^t \lambda_{02}(v)\,dv\right)$$

Assumptions:

- i.i.d. observations
- independent censoring
- independence of the stochastic processes 0→1 and 0→2

Survival with ‚chemo-therapy only'

Assumption:
independence of the stochastic processes 0→1 and 0→2

$\hat{S}_0(t)$ **KME with transitions to state 1 censored.**

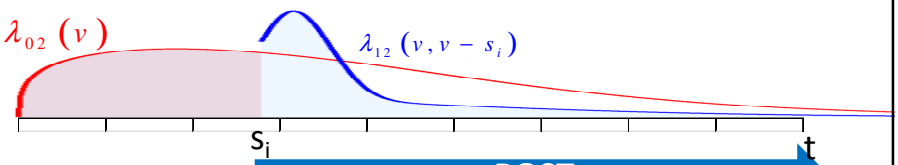
Survival with chemo-therapy

$$S_0(t) = \exp\left(-\int_0^t \lambda_{02}(v)\,dv\right)$$

n pseudo-values for survival in state 0 based on $S_0(t)$

Transition at time s_i

Survival after s_i: $S(t|t_0>s_i)$
Pseudo-value based on $\hat{S}(t|t_0 = s_i)$

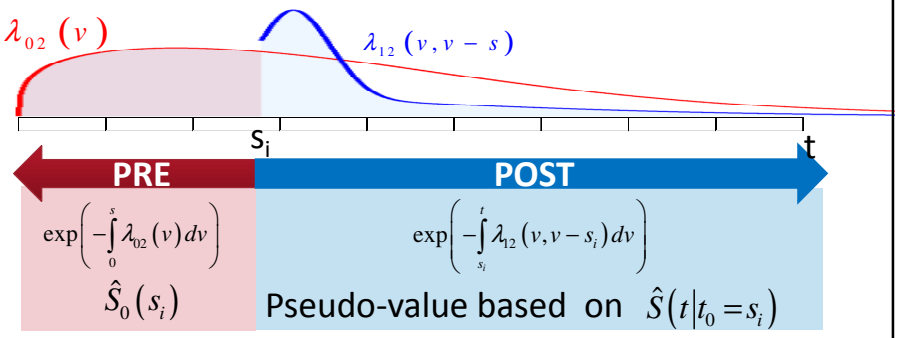


POST

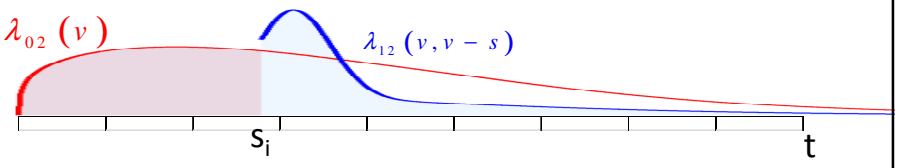
Survival when SCT is done at time s_i

$$\exp\left(-\int_{s_i}^t \lambda_{12}(v, v-s_i)\,dv\right)$$

Transition at time si



For each observed s_i



$$\exp\left(-\left(\int_0^s \lambda_{02}(v)dv + \int_s^t \lambda_{12}(v, v-s)dv\right)\right)$$

PRE- AND POST RISK

n_s modified pseudo-values for survival with transition at times s_i

- **n pseudo-values for survival in state 0 based on S_0**

$$E\left(\hat{V}_{i,0}(x)\right) = S_0(x) = \exp\left(-\int_0^x \lambda_{02}(v)\,dv\right)$$

Survival, when state 1 is eliminated (i.e. no SCTs are performed)

- **n_s modified pseudo-values for survival with transition at times s_i**

$$E\left(\hat{V}_{i,1}\left(x \middle| t_{01} = s_i\right)\right) = \exp\left(-\left(\int_0^{s_i} \lambda_{02}(v)\,dv + \int_{s_i}^{t} \lambda_{12}(v, v-s)\,dv\right)\right)$$

Intention to treat perspective, when transition time is s_i

Generalised linear Model

- **Response**:
  n pseudo-values for survival in state 0 based on S_0
  n_s modified pseudo-values for survival with transition at times s_i → $n+n_s$ pseudovalues
- **Response probability distribution**:
  Normal distribution
- **Independent variables**
  Indikator variable + (>=0) factors + intercept
- **Link function g():**
  e.g. log(-log()), logit(), ...
- **Variance estimation:**
  empirical sandwich estimator;

Generalised linear Model

With one explanatory variable

- **Intercept β_0: Survival without SCT**

$$\hat{\beta}_0 = g(\bar{V}_0) \qquad E(\bar{V}_0) = S_0(t) = \exp\left(-\int_0^t \lambda_{02}(v)\,dv\right)$$

- β_1 : $\qquad \hat{\beta}_1 = g(\bar{V}_1) - g(\bar{V}_0)$
- **$\beta_0 + \beta_1$: Expected survival with SCT, conditional on waiting times in a given sample**

$$\hat{\beta}_0 + \hat{\beta}_1 = g(\bar{V}_1) \qquad E(\bar{V}_1) = S_1(t) = \int_0^t \left(q(s) \exp\left(-\left(\int_0^s \lambda_{02}(v)\,dv + \int_s^t \lambda_{12}(v, v-s)\,dv \right) \right) \right) ds$$

q(s) is distribution of waiting times in a given sample

p=1 $\hat{\beta}_0 = g(\bar{V}_0) \qquad E(\bar{V}_0) = S_0(t) = \exp\left(-\int_0^t \lambda_{02}(v)\,dv\right)$

$$\hat{\beta}_0 + \hat{\beta}_1 = g(\bar{V}_1) \qquad E(\bar{V}_1) = S_1(t) = \int_0^t \left(q(s) \exp\left(-\left(\int_0^s \lambda_{02}(v)\,dv + \int_s^t \lambda_{12}(v, v-s)\,dv \right) \right) \right) ds$$

q(s) is distribution of waiting times in a given sample

$$\hat{\beta}_1 = g(\bar{V}_1) - g(\bar{V}_0)$$

Log-log link

$$\exp(\beta_1) = \left(\frac{\log(S_1(t))}{\log(S_0(t))} \right)$$

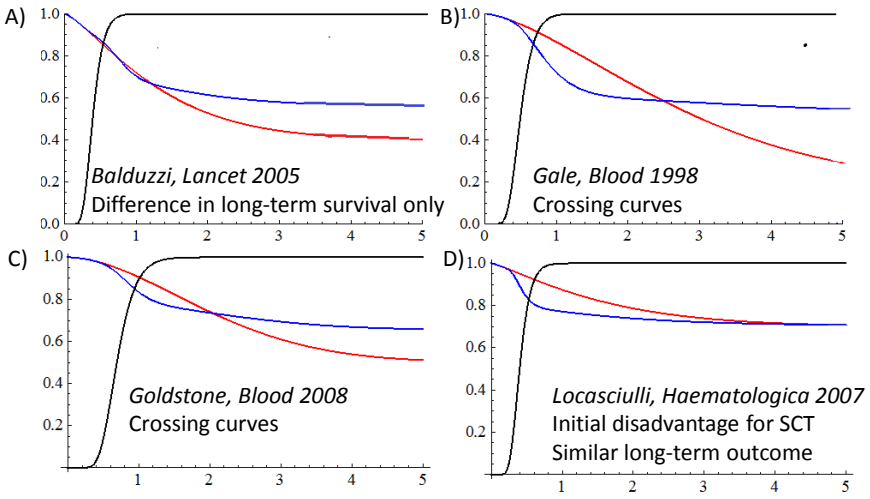
Response: $\hat{\mathbf{V}} = \left(\hat{V}_{1,0}(t),...,\hat{V}_{n,0}(t),\hat{V}_{1,1}\left(t\middle|t_{01}=s_1\right),...,\hat{V}_{n_s,1}\left(t\middle|t_{01}=s_{n_s}\right)\right)'$

Independent factors: $\mathbf{X} = \left(\mathbf{X}_0,...,\mathbf{X}_p\right)'$

X_1 : Indicator variable indicating whether a transition to State 1 occured

$$g\left(E\left(\hat{\mathbf{V}}\right)\right) = \mathbf{X}'\boldsymbol{\beta}$$

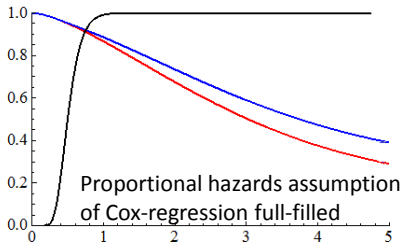
Simulation study: Scenarios



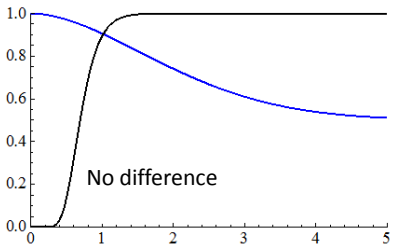
Without SCT/Donor, with SCT/Donor, observed waiting time distribution

Simulation study: Scenarios 2

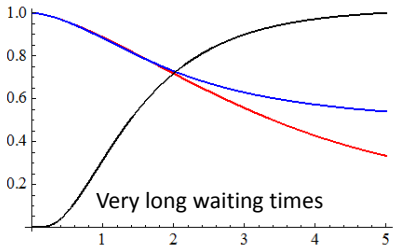
E)



F)



G)



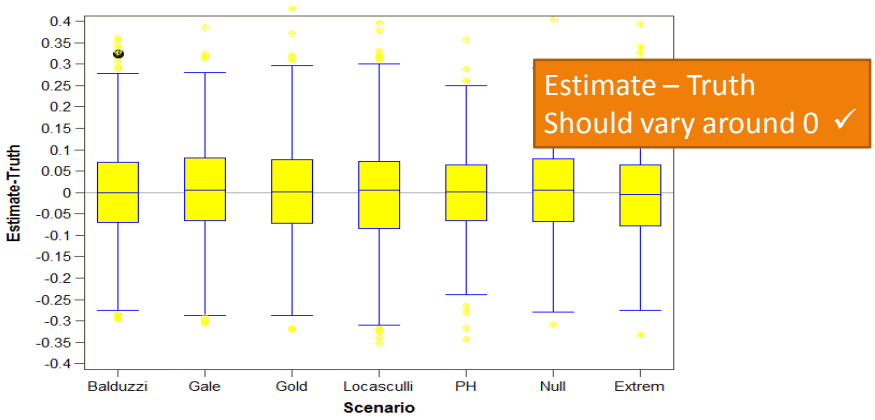
Without SCT/Donor, with SCT/Donor, observed waiting time distribution

Truths

Scenario		TRUTH		ITT
	based on	In State 0 (no SCT)[1]	With transition 01 (SCT)[2]	Susceptible (with Donor)[3]
A	Balduzzi	40%	56%	56%
B	Gale*	29%	55%	55%
C	Goldstone	51%	66%	66%
D	Loscatiulli*	71%	71%	71%
E	PH	29%	39%	39%
F	No diff.	51%	51%	51%
G	Late SCTs	33%	61%	54%

log HR based on 5-years pSU

1000 simulations a 1000 patients



Estimate – Truth
Should vary around 0 ✓

Coverage of novel approach

Scenario		Coverage					
		Below		within		Above	
	based on	n	%	n	%	n	%
A	Balduzzi	23	2.3%	952	95.2%	25	2.5%
B	Gale	30	3.0%	942	94.2%	28	2.8%
C	Goldstone	25	2.5%	958	95.8%	17	1.7%
D	Loscatiulli	26	2.6%	949	94.9%	25	2.5%
				956	95.6%	25	2.5%
				956	95.6%	18	1.8%
				957	95.7%	17	1.7%

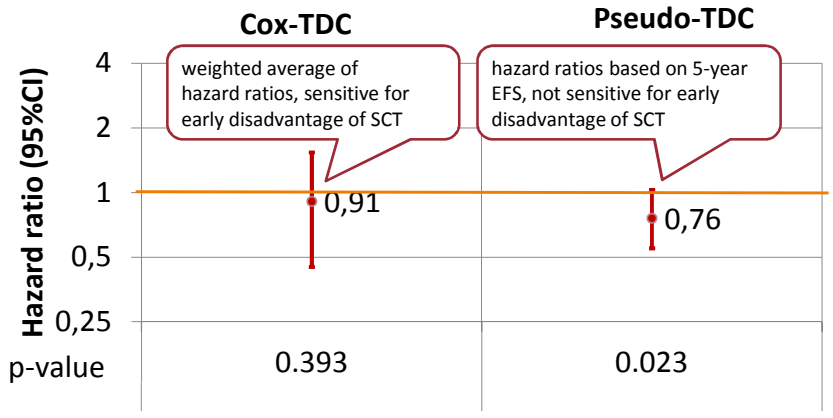
Coverage
In a simulation study ~ 95% of the simulated 95% confidence-intervals need to include the the true value
if the new method works correctly

Should vary around 95% ✓

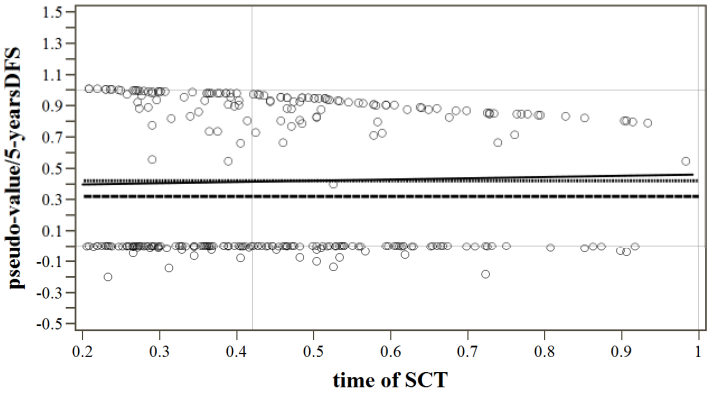
Similar results

- For β_0
- **For n = 400**

Motivating Example continued



Impact of time of SCT



$\hat{S}_1(5) = 0.42 \pm 0.03$

$\hat{S}_0(5) = 0.32 \pm 0.03$

pseudo-values with SCT and regression line

5-yrs DFS Chemotherapy

5-yrs DFS SCT (mean)

Properties of novel approach

Problem	Cox	Pseudo	LM	novel
PH-assumption not needed	×	✓	?	✓
addresses long-term outcome	×	✓	✓	✓
adjustment for waiting-time bias	✓	×	✓	✓
no arbitrary specification landmark-time needed	✓	×	×	✓
inclusion of covariates	✓	✓	×	✓
parameter estimates clear interpretation	✓	✓	×	✓

But needs to choose a time-point for comparison of long-term survival

Conclusion

- Methods for long-term outcome are useful in pediatric oncology
- if non-proportional hazards occur in combination with time-dependent covariates, novel approaches are needed
- the proposed approach has a clear interpretation and directly answers the question of long-term outcome

Introduction Pseudo-values SCT-Chemo Pseudo-TDC Conclusion

Conclusion

- **Physician:** ‚Should I perform an SCT in PH+ALL?
- **Answer:** Yes, expected 5-years DFS without SCT is 32%, and expected 5-years DFS with SCT is 42%

IBS-ROeS Seminar 2013
September 9th-12th, 2013 – Dornbirn, Austria

Thanks

- M. Mittlböck, PhD and
  H. Heinzl, PhD
  Section for Clinical Biometrics
  Medical University of Vienna
- MG Valsecchi, PhD
  University of Milano-Bicocca

Thanks for your attention