Competing risks models and time-dependent effects

Hein Putter

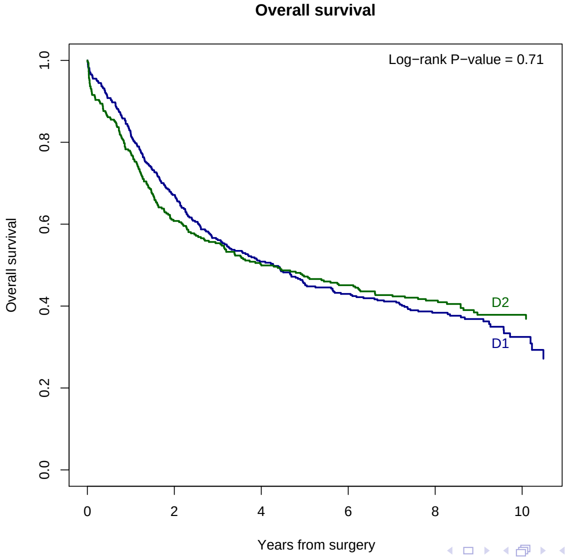
Department of Medical Statistics and Bioinformatics
Leiden University Medical Center

IBS British/Irish Region and RSS Medical Section
October 7, 2008

Dutch Gastric Cancer Trial

- Randomized clinical trial for gastric cancer patients
- 711 patients randomized between D1-dissection (380) and D2-dissection (331)
- D1-dissection: limited lymph-node dissection
  - Standard in Western Europe and America
- D2-dissection: more extensive lymph-node dissection
  - Standard in Japan
  - More surgical complications, more hospital deaths
  - Better overall survival in the lung run?
- Median follow-up: 9.1 years

Kaplan-Meier survival curves



Time-fixed Cox regression

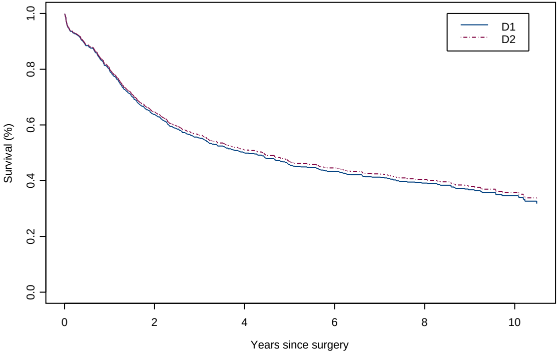
- Suppose single covariate Z, say treatment
- $Z = 0$, placebo or standard treatment, here D1
- $Z = 1$, experimental treatment, here D2
- Hazard rate given by

$$\lambda(t \,|\, Z) = \lambda_0(t) \exp(\beta Z)$$

- $\lambda_0(t)$ is hazard rate corresponding to D1, $\lambda_1(t)$ to D2
- Proportional hazards assumption: ratio of hazards $\lambda_1(t)/\lambda_0(t) = \exp(\beta)$, independent of t

Applied to the Dutch Gastric Cancer Trial

- Result: $\hat{\beta} = -0.033$, se = 0.095, hazard ratio = 0.97, $P = 0.73$



Time-dependent Cox regression

- Hazard rate given by

$$\lambda(t \mid Z) = \lambda_0(t) \exp(\beta_F Z + \beta_T Z f(t))$$

- β_F: time-fixed regression coefficient
- β_T: time-dependent regression coefficient
- Hazard ratio now varies with t

$$HR(t) = \frac{\lambda_1(t)}{\lambda_0(t)} = \exp(\beta_F + \beta_T f(t))$$

- Can test hypothesis $\beta_T = 0$; $\beta_T \neq 0$ indicates time-dependent treatment effect, violation of proportional hazards assumption

Applied to the Dutch Gastric Cancer Trial

- Take $f(t) = \log(t+1)$:

	coefficient	standard error	P-value
treatment	0.760	0.261	0.0036
time*treatment	-0.273	0.084	0.0011

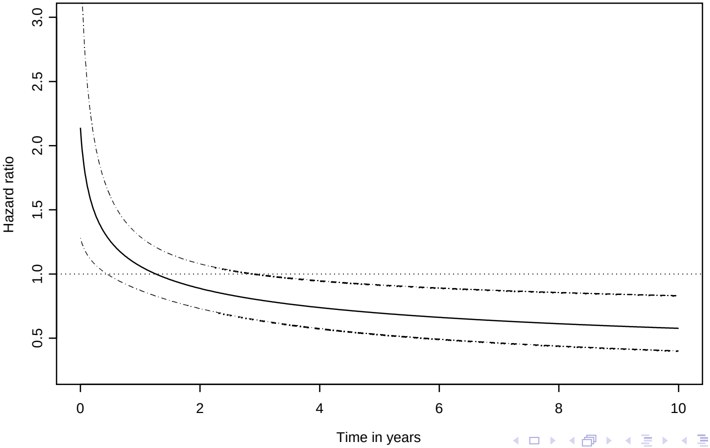
- Significant interaction of time by treatment;
- Indicates that proportional hazards assumption is violated;
- "Hazard ratio" (HR) will vary over time, according to

$$HR(t) = \exp(0.760 - 0.273 * \log(t+1)) \; ;$$

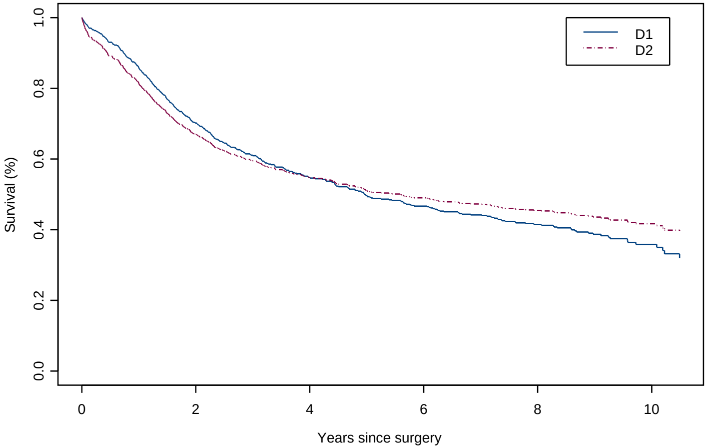
- $\mathrm{se}(\log(HR(t))) = \sqrt{\mathrm{var}(\hat\beta_F) + \mathrm{var}(\hat\beta_T) f^2(t) + 2\mathrm{cov}(\hat\beta_F, \hat\beta_T) f(t)}$.

Hazard ratio over time

Hazard ratio of D2 with respect to D1:



Fitted



Competing risks point of view

- So far we have looked at overall survival
- Cause of death also recorded

	Gastric cancer	Other causes
D1 ($n = 380$)	178 (47 %)	74 (19 %)
D2 ($n = 331$)	121 (37 %)	72 (22 %)

Competing risks

- Death due to different causes (competing risks)
- We observe
  - Time of death T
  - Cause of death D (1=cancer, 2=other, 0=censoring)
- Death of one cause precludes death of other cause
- Fundamental concept in competing risks models: the cause-specific hazard function

$$\lambda_k(t) = \lim_{\Delta t \downarrow 0} \frac{P(t \leq T < t + \Delta t, D = k \mid T \geq t)}{\Delta t}$$

- Cumulative cause-specific hazard

$$\Lambda_k(t) = \int_0^t \lambda_k(s) ds$$

Total and relative hazards

- Total hazard

$$\lambda_\bullet(t) = \sum_{k=1}^{K} \lambda_k(t)$$

- Relative hazard

$$\pi_k(t) = \frac{\lambda_k(t)}{\lambda_\bullet(t)}$$

Total and relative hazards

- Total hazard

$$\lambda_\bullet(t) = \sum_{k=1}^{K} \lambda_k(t)$$

- Relative hazard

$$\pi_k(t) = \frac{\lambda_k(t)}{\lambda_\bullet(t)}$$

- Total hazard drives overall survival

$$S(t) = \exp(-\Lambda_\bullet(t))$$

Total and relative hazards

- Total hazard

$$\lambda_\bullet(t) = \sum_{k=1}^{K} \lambda_k(t)$$

- Relative hazard

$$\pi_k(t) = \frac{\lambda_k(t)}{\lambda_\bullet(t)}$$

- Total hazard drives overall survival

$$S(t) = \exp(-\Lambda_\bullet(t))$$

- Relative hazard drives cause of death (given death at t)

Cumulative incidence functions

- Cumulative incidence function, $I_k(t)$, is defined as the probability of failing from cause k before time t

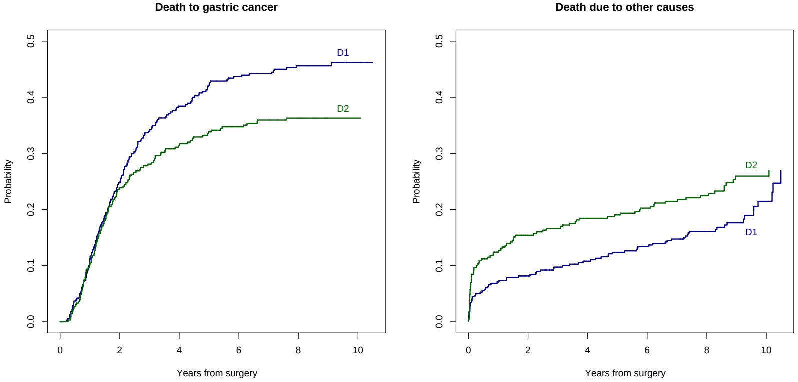
$$I_k(t) = P(T \leq t, D = k) = \int_0^t \lambda_k(s) S(s) ds$$

- Estimation of $I_k(t)$ by replacing hazards by estimates in above formula
- Don't use naive Kaplan-Meier estimator; is estimating

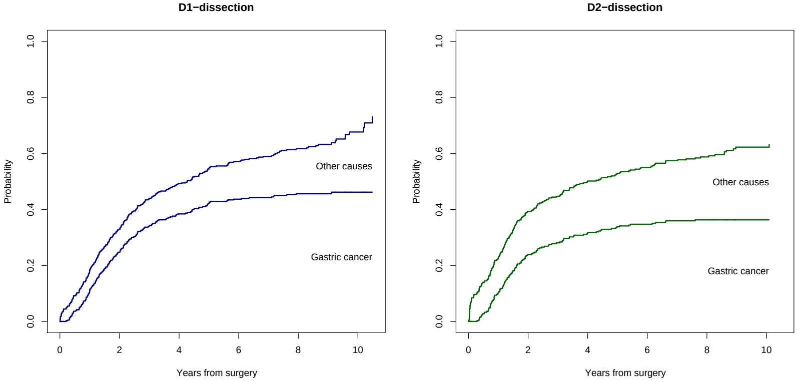
$$1 - S_k(t) = \int_0^t \lambda_k(s) S_k(s) ds \ ,$$

rather than $I_k(t)$

Cumulative incidence curves



Organized differently (stacked)



Regression on cause-specific hazards

- $\lambda_k(t \mid Z)$: cause-specific hazard of cause k of a subject with treatment Z (with Z=0 (D1) or 1 (D2)) given by

$$\lambda_k(t \mid Z) = \lambda_{k,0}(t) \exp(\beta_k Z) \ ,$$

  with
  - $\lambda_{k,0}(t)$ is the baseline (D1) cause-specific hazard of cause k
  - β_k represents the treatment effect on cause k
- Analysis completely standard
- But effect of covariate on cumulative incidence sometimes unexpected (because it depends on the competing hazards as well)!

Regression on cumulative incidence function

- Define subdistribution hazard

$$\overline{\lambda}_k(t) = -\frac{d \log(1 - I_k(t))}{dt} .$$

- Impose proportional hazards assumption on the subdistribution hazard instead of on cause-specific hazard [Fine & Gray, 1999]

$$\overline{\lambda}_k(t \,|\, Z) = \overline{\lambda}_{k,0}(t) \exp(\beta_k Z) .$$

- This does imply expected relation

Regression results

	Death due to gastric cancer		Death due to other causes	
	Coef (SE)	P-value	Coef (SE)	P-value
Cause-specific hazards	-0.18 (0.12)	0.12	0.25 (0.16)	0.12
Fine-Gray	-0.27 (0.12)	0.021	0.31 (0.16)	0.055
Gray log-rank		0.021		0.061

- Cause-specific hazard ratios of D2 with respect to D1
  - Gastric cancer: $HR_1 = 0.83$
  - Other causes: $HR_2 = 1.28$

Back to overall survival

- Question: can a proportional cause-specific hazards model explain time-dependent effects seen on overall survival?

Back to overall survival

- Question: can a proportional cause-specific hazards model explain time-dependent effects seen on overall survival?
- Total hazard:
  - D1-dissection: $\lambda_{\bullet,D1}(t) = \lambda_{1,0}(t) + \lambda_{2,0}(t)$
  - D2-dissection: $\lambda_{\bullet,D2}(t) = \lambda_{1,0}(t) HR_1 + \lambda_{2,0}(t) HR_2$
- Relative hazard for D1:
  $\pi_{1,D1}(t) = \frac{\lambda_{1,0}(t)}{\lambda_{1,0}(t)+\lambda_{2,0}(t)} = 1 - \pi_{2,D1}(t)$

Back to overall survival

- Question: can a proportional cause-specific hazards model explain time-dependent effects seen on overall survival?
- Total hazard:
  - D1-dissection: $\lambda_{\bullet,D1}(t) = \lambda_{1,0}(t) + \lambda_{2,0}(t)$
  - D2-dissection: $\lambda_{\bullet,D2}(t) = \lambda_{1,0}(t) HR_1 + \lambda_{2,0}(t) HR_2$
- Relative hazard for D1:
  $\pi_{1,D1}(t) = \frac{\lambda_{1,0}(t)}{\lambda_{1,0}(t)+\lambda_{2,0}(t)} = 1 - \pi_{2,D1}(t)$
- Hazard ratio D2 with respect to D1 for overall survival

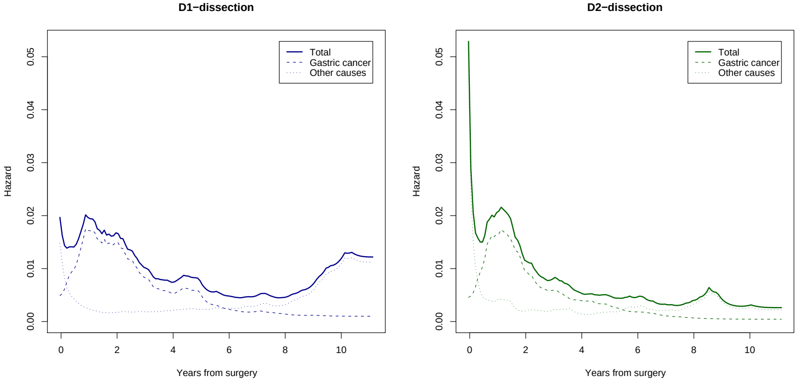
$$
\begin{aligned}
HR(t) &= \frac{\lambda_{\bullet,D2}(t)}{\lambda_{\bullet,D1}(t)} = \frac{\lambda_{1,0}(t) HR_1 + \lambda_{2,0}(t) HR_2}{\lambda_{1,0}(t) + \lambda_{2,0}(t)} \\
&= \pi_{1,D1}(t) HR_1 + (1 - \pi_{1,D1}(t)) HR_2
\end{aligned}
$$

Back to overall survival

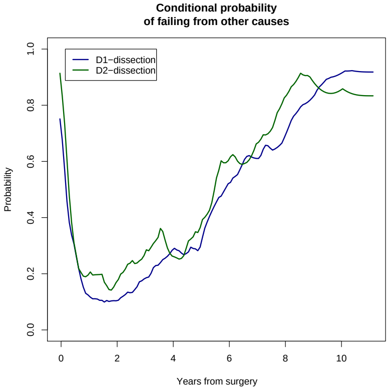
- Question: can a proportional cause-specific hazards model explain time-dependent effects seen on overall survival?
- Total hazard:
  - D1-dissection: $\lambda_{\bullet,D1}(t) = \lambda_{1,0}(t) + \lambda_{2,0}(t)$
  - D2-dissection: $\lambda_{\bullet,D2}(t) = \lambda_{1,0}(t)HR_1 + \lambda_{2,0}(t)HR_2$
- Relative hazard for D1:
  $\pi_{1,D1}(t) = \frac{\lambda_{1,0}(t)}{\lambda_{1,0}(t)+\lambda_{2,0}(t)} = 1 - \pi_{2,D1}(t)$
- Hazard ratio D2 with respect to D1 for overall survival

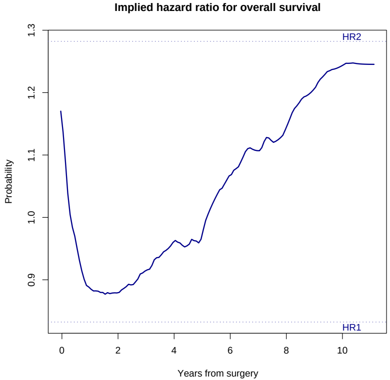
$$
\begin{aligned}
HR(t) &= \frac{\lambda_{\bullet,D2}(t)}{\lambda_{\bullet,D1}(t)} = \frac{\lambda_{1,0}(t)HR_1 + \lambda_{2,0}(t)HR_2}{\lambda_{1,0}(t) + \lambda_{2,0}(t)} \\
&= \pi_{1,D1}(t)HR_1 + (1 - \pi_{1,D1}(t))HR_2 \\
\text{(or)} &= \pi_{2,D1}(t)HR_2 + (1 - \pi_{2,D1}(t))HR_1
\end{aligned}
$$

Cause-specific and total hazards (smoothed)



Relative hazards





Modeling approaches

- Let's see if we can use the concepts of total and relative hazards in a model
- Regression on cause-specific hazards models effect of treatment for each cause separately
- This is an indirect, but convenient way of modeling the joint distribution, $P(T, D)$, of time and cause of failure (recall, $I_k(t) = P(T \leq t, D = k)$)
- Other approaches
  - $P(T, D) = P(D) \cdot P(T \mid D)$ [Larson & Dinse, 1985], mixture model

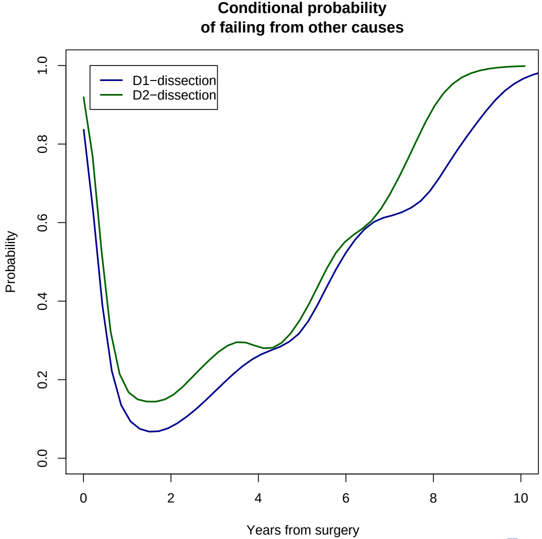
Modeling approaches

- Let's see if we can use the concepts of total and relative hazards in a model
- Regression on cause-specific hazards models effect of treatment for each cause separately
- This is an indirect, but convenient way of modeling the joint distribution, $P(T, D)$, of time and cause of failure (recall, $I_k(t) = P(T \leq t, D = k)$)
- Other approaches
  - $P(T, D) = P(D) \cdot P(T \mid D)$ [Larson & Dinse, 1985], mixture model
  - $P(T, D) = P(T) \cdot P(D \mid T)$, selection model, cause of failure model

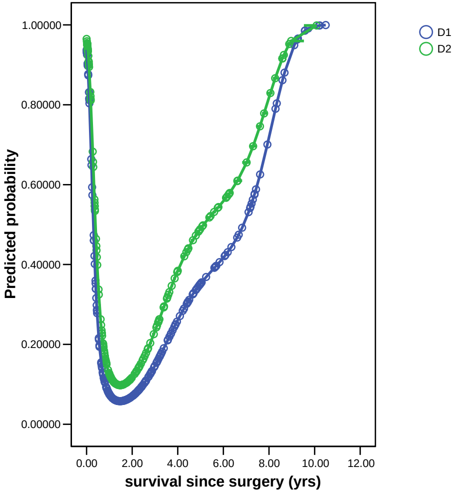
Cause of failure model

- Model for $P(T)$: for instance (but not necessarily) proportional hazards model
- Model for $P(D \mid T = t)$: logistic (general polytomous) regression with treatment and time-effects
- Two approaches:
  - Logistic regression model for D1- and D2-dissection separately, with smoothed time-effects
  - Logistic regression model with treatment, $f(t)$, and interaction

Smoothed directly with logistic regression



With SPSS (linear, quadratic plus cubic)



Cumulative incidence functions

- How do we get the cumulative incidence functions back?
- Remember we have model for total hazard, and hence we have an estimate of $\exp(-\Lambda_\bullet(t))$
- Now we have

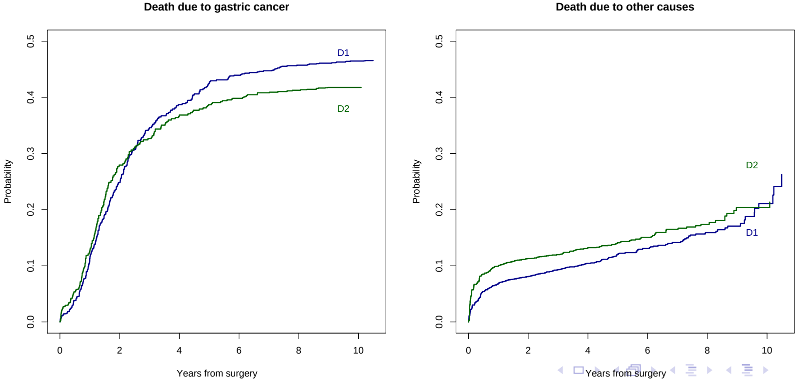
$$\lambda_1(t) = \pi_1(t)\,\lambda_\bullet(t)$$

- Hence

$$\begin{aligned} I_1(t) &= \int_0^t \lambda_1(s)\,\exp(-\Lambda_\bullet(s))ds \\ &= \int_0^t \pi_1(s)\,\lambda_\bullet(s)\,\exp(-\Lambda_\bullet(s))ds \\ &= \int_0^t \pi_1(s)\,f_\bullet(s)ds \end{aligned}$$

Resulting cumulative incidence curves

- Based on
  - Separate logistic regression models for D1 and D2 with smoothed time effects
  - Separate Kaplan-Meier estimates for D1 and D2



Discussion

- Competing risks with different cause-specific treatment effects can explain time-dependent effects
  - Overall hazard ratio is weighted mean of cause-specific hazard ratios, weights given by relative hazards
- Mixture model:
  - Suggests that cause of failure is determined from the outset
  - Typically need EM-algorithm for estimation
  - Results depend on follow-up
- Selection model, cause of failure model:
  - Useful alternative to proportional cause-specific or subdistribution hazards model
  - Straightforward to program, but may need smoothing of $P(D = d \mid T = t)$ as function of t
  - Can incorporate missing cause of failure

References

J. P. Fine and R. J. Gray.
A proportional hazards model for the subdistribution of a competing risk.
J Am Stat Ass **94**:496–509, 1999.

M. G. Larson, and G. E. Dinse.
A mixture model for the regression analysis of competing risks data.
Appl Stat **34**: 201–211, 1985.

H. Putter, M. Sasako, H. H. Hartgrink, C. J. H. van de Velde, and J. C. van Houwelingen.
Long-term survival with non-proportional hazards: results from the Dutch Gastric Cancer Trial.
Stat Med **24**: 2807–2821, 2005.

H. Putter, M. Fiocco, R. B. Geskus.
Tutorial in biostatistics: Competing risks and multi-state models.
Stat Med **26**: 2389–2430, 2007.