Three Equivalent Representations of the Cause-Specific Cumulative Incidence Estimator and the Fine & Gray Model

The Fine & Gray model for competing risks with left truncated data or time-dependent covariates

Ronald Geskus

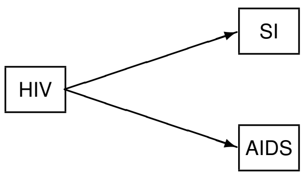
Amsterdam Health Service; Academic Medical Center

October 7, 2008

Motivating Example: HIV/AIDS

- Since 1995, powerful anti-HIV treatment (HAART) available
- AIDS incidence greatly reduced
- More individuals remain alive
- Increase in non-HIV related death on cumulative scale after 1995?

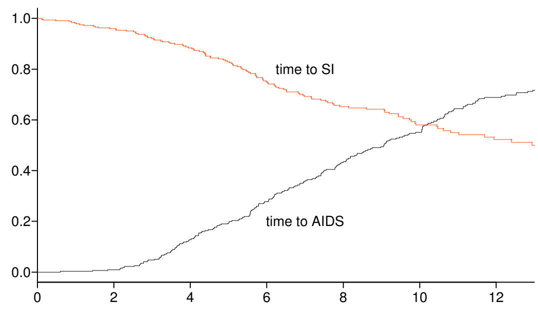
Example Competing Risks



- Time from HIV infection to switch to syncytium-inducing phenotype (SI) or AIDS as first event
- Outcomes may be dependent:
  both SI switch and AIDS more likely at advanced progression
- Effect of genetic cofactor "CCR5-Δ32 deletion": WW and WM

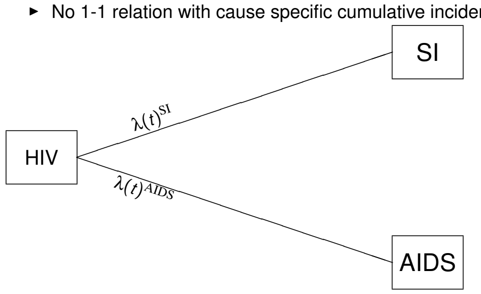
Three types of hazards

1. **marginal hazard**
   only estimable under independence (⟶ Kaplan-Meier)

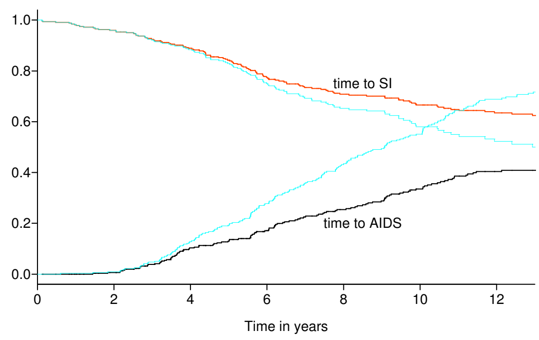


Three types of hazards

1. marginal hazard
   only estimable under independence (⟶ Kaplan-Meier)
2. **cause-specific transition hazard**
   - Censor individuals with competing event
   - Standard Cox analysis, but interpretation different
   - No 1-1 relation with cause specific cumulative incidence $P^{\mathrm{SI}}(0,t)$



Cause specific cumulative incidence



$P^{\mathrm{SI}}(0,10) \approx 0.35$ and $P^{\mathrm{AIDS}}(0,10) \approx 0.35$

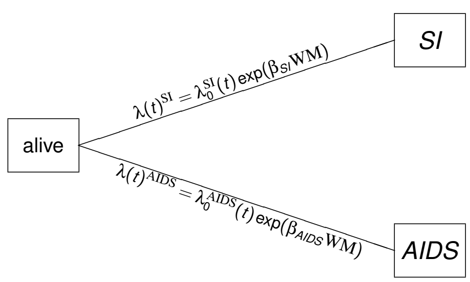
Three types of hazards

1. marginal hazard
   only estimable under independence (⟶ Kaplan-Meier)
2. cause-specific transition hazard
   - Censor individuals with competing event
   - Standard Cox analysis, but interpretation different
   - No 1-1 relation with cause specific cumulative incidence $P^{\mathrm{SI}}(0,t)$
3. **subdistribution** (cause specific cumulative incidence)

$$\lambda^{*\mathrm{SI}}(t) = -\frac{d\log(1 - P^{\mathrm{SI}}(0,t))}{dt}$$

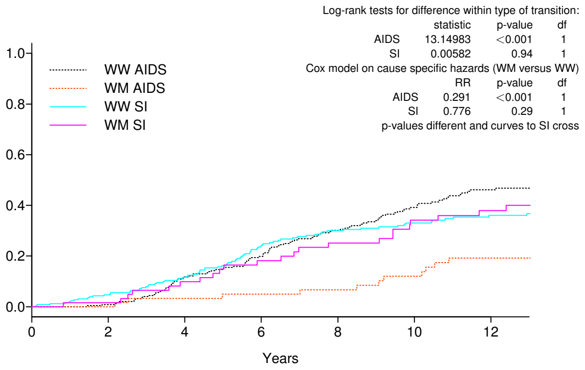
4. **Fine & Gray model**: proportional hazards regression on λ^*
   They did not consider left truncated data and time-dependent covariates

Regression on Cause Specific Hazard

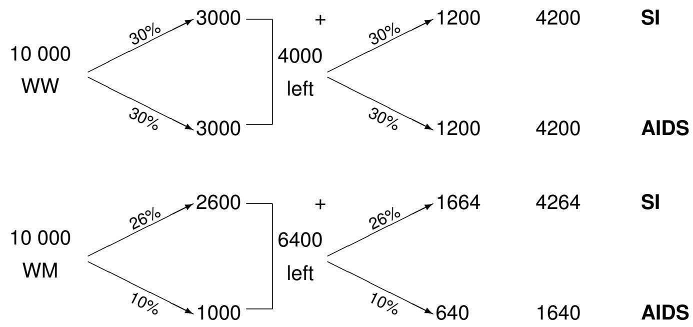


	coef	exp(coef)	se(coef)	p
β_{SI}	-0.254	0.776	0.238	0.290000
β_{AIDS}	-1.236	0.291	0.307	0.000057

Cumulative progression to SI/AIDS by CCR5



"Reproducing" this result: CSH-RR $\neq$ cumulative effect



$\text{CSH} - \text{RR(AIDS)} = 10/30 = 0.33$
$\text{CSH} - \text{RR(SI)} = 26/30 = 0.87$

Results effect of CCR5-Δ32 on progression to SI

	F&G	CS	log-rank
coefficients:	**0.0236**	-0.254	
standard errors:	0.227	0.238	
p-values:	**0.92**	0.29	**0.94**

Right Censored Data

Observed data $(x_1, \delta_1), \ldots, (x_N, \delta_N)$; $x_i = t_i \wedge c_i$, $\delta_i = \{t_i \le c_i\}$
$t_{(i)}$ are ordered **unique** event times
$r(t_{(i)})$ number at risk, $d(t_{(i)})$ number of events at $t_{(i)}$
m_j **number of censorings at** $c_{(j)}$

Kaplan-Meier: $1 - \widehat{F}(t) = \widehat{\bar{F}}^{\text{PL}}(t) = \prod_{i:t_{(i)} \le t} \left(1 - \frac{d(t_{(i)})}{r(t_{(i)})}\right)$

- Satten & Datta (Am. Stat.; 2001):

$$\widehat{F}^{\text{PL}}(t) \equiv \widehat{F}^{\text{IPCW}}(t) := \frac{1}{N} \sum_{i:t_{(i)} \le t} \frac{d(t_{(i)})}{\widehat{\bar{G}}(t_{(i)}-)}$$

if

$$\widehat{\bar{G}}(t-) = \prod_{j:c_{(j)} < t} \left[1 - \frac{m_j}{r(c_{(j)}) - d(c_{(j)})}\right].$$

Left Truncated Data

Observed data $(l_1, t_1), \ldots, (l_n, t_n)$
l_i entry (left truncation) time; $H(l) = P(L \le l)$

Kaplan-Meier: $1 - \widehat{F}(t) = \widehat{\bar{F}}^{\text{PL}}(t) = \prod_{i:t_{(i)} \le t} \left(1 - \frac{d(t_{(i)})}{r(t_{(i)})}\right)$

- Shen (Comm. in Stat. 2003):

$$\widehat{F}^{\text{PL}}(t) \equiv \widehat{F}^{IPTW}(t) := \frac{1}{\sum_{i=1}^{n} \frac{d(t_{(i)})}{\widehat{H}(t_{(i)})}} \times \sum_{i:t_{(i)} \le t} \frac{d(t_{(i)})}{\widehat{H}(t_{(i)})}$$

if

$$\widehat{H}(t) = \prod_{l_{(j)} > t} \left[1 - \frac{w_j}{r(l_{(j)})}\right]$$

- Note $\sum_{i=1}^{n} \frac{d(t_{(i)})}{\widehat{H}(t_{(i)})}$ is estimator of $\frac{n}{P(L \le T)}$

Same equivalence with both left truncated and right censored data

$$\widehat{F}^{IPW}(t) = \frac{1}{\sum_{i=1}^{n} \frac{d(t_{(i)})}{\widehat{H}(t_{(i)})} + \sum_{j=1}^{m} \frac{m_j}{\widehat{H}(c_{(j)})}} \times \sum_{i:t_{(i)} \le t} \frac{d(t_{(i)})}{\widehat{\bar{G}}(t_{(i)}-) \times \widehat{H}(t_{(i)})}$$

Basic idea of proof:

- Similar to Shen, first consider $X = T \wedge C$
- L is right truncated by $X = T \wedge C$
- Use special product limit form of $\widehat{G}$ and $\widehat{F}^{PL}$

$$P(X \ge t_{(j)}) = \left[1 - \widehat{G}(t_{(j)}-)\right] \times \left[1 - \widehat{F}^{PL}(t_{(j)}-)\right]$$

can be extended to competing risks setting

Standard and IPW Estimator

- Observed data $\{(l_1, x_1, d_1\delta_1), \ldots, (l_N, x_N, d_N\delta_N)\}$
- Standard estimator cause-specific cumulative incidence

$$\widehat{F}_k(t) = \sum_{i:t_{(i)} \le t} \widehat{\lambda}_k(t_{(i)}) \widehat{\bar{F}}(t_{(i)}-)$$

- Equal to IPW estimator

$$\widehat{F}_k^{IPW}(t) = \frac{1}{\sum_{i=1}^{n} \frac{d(t_{(i)})}{\widehat{H}(t_{(i)})} + \sum_{j=1}^{m} \frac{m_j}{\widehat{H}(c_{(j)})}} \times \sum_{i:t_{(i)} \le t} \frac{d_k(t_{(i)})}{\widehat{\bar{G}}(t_{(i)}-) \times \widehat{H}(t_{(i)})}$$

Product Limit Estimator

- Cause specific cumulative incidence F_k described by

$$T^* = T \times I\{D = k\} + \infty \times I\{D \neq k\}$$

- Without censoring/truncation, individuals with competing event remain in risk set forever
  Hence, standard software can be used with small change in data
- Subdistribution hazard of T^*

$$\lambda_k^*(t) = \frac{f_k^*(t)}{\bar{F}(t-) + \sum_{s \neq k} F_s(t-)}$$

- **Standard estimator equal to** $\widehat{\bar{F}}_k^{PL}(t) = \prod_{i:t_{(i)} \le t} \left[1 - \widehat{\lambda}_k^*(t_{(i)})\right]$

Left Truncated/Right Censored Data

- $\widehat{\lambda}^*_k(t_{(i)}) = \dfrac{d_k(t_{(i)})}{r^*(t_{(i)})}$ with r^* virtual number at risk
- Contribution to $r^*(t_{(i)}) = r(t_{(i)}) + \sum \omega_j(t_{(i)})$
  - censored before $t_{(i)}$: 0
  - still at risk at $t_{(i)}$: 1
  - competing event at t_j before $t_{(i)}$: weight

$$\omega_j(t_{(i)}) = \frac{\widehat{G}(t_{(i)}-)}{\widehat{G}(t_j-)} \times \frac{\widehat{H}(t_{(i)})}{\widehat{H}(t_j)}$$

Calculation

- Standard software can be used, with weighted data set

id	Tstart	Tstop	stat
1	0.25460	0.63644	0
2	0.00000	0.64358	1
3	0.08005	0.25615	2

$\Longrightarrow$

id	Tstart	Tstop	stat	weight.cens	weight.trunc
1	0.25460	0.63644	0	1.00000	1.00000
2	0.00000	0.64358	1	1.00000	1.00000
3	0.08005	0.25615	2	1.00000	1.00000
3	0.25615	0.31778	2	1.00000	1.02941
3	0.31778	0.37693	2	1.00000	1.02941
3	0.37693	0.38928	2	1.00000	1.02941
3	0.38928	0.46029	2	1.00000	1.02941
3	0.46029	0.50979	2	1.00000	1.02941
3	0.50979	0.64358	2	0.67849	1.07230
3	0.64358	0.64724	2	0.67849	1.07230

Events of type 1 observed at 0.31778 0.37693 0.38928 0.46029 0.50979 0.64358 0.64724

- In R:
```
survfit(Surv(Tstart,Tstop,stat==1),
        data=wData,weight=weight.cens*weight.trunc)
```

The Fine & Gray Model

- Cox regression on subdistribution hazard
- Solution to ω-weighted score function of partial likelihood
  Implemented in `cmprsk` R package
- Alternative: use weighted data set with standard Cox partial likelihood software
  In R:
```
coxph(Surv(Tstart,Tstop,stat==1),...,
      weight=weight.cens*weight.trunc)
```
- Estimator for standard error:
  - Fine & Gray use sandwich type estimator in R: use `cluster(id)`
  - What about standard estimator?

Simulation Study: Right Censored Data

- Two covariates, Z_1 and $Z_2 \sim N(0,1)$, effect $\beta_1 = \beta_2 = 0.5$
- T_1^* unit exponential with mass 0.7 at ∞ if $Z_1 = Z_2 = 0$
- Censoring uniform on $[0.5, 1]$

Results for β_1 (5000 runs):

	N	bias	p-value	$\widehat{\text{var}(\beta_1)}$	$\overline{\text{var}(\beta_1)}$	$P(0.5 \in \text{CI})$	p-value
`coxph`	100	0.018	$<$0.001	0.069	0.065	0.946	0.21
	200	0.005	0.043	0.032	0.030	0.944	0.074
	500	0.001	0.38	0.012	0.011	0.950	1.00
`crr` or	100	0.018	$<$0.001	0.069	0.062	**0.928**	$<$0.001
`cluster(id)`	200	0.005	0.043	0.032	0.029	**0.934**	$<$0.001
	500	0.001	0.38	0.012	0.011	0.945	0.14

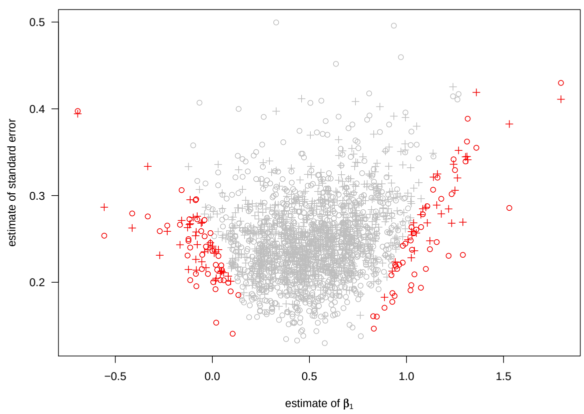
Simulation Study: Left Truncated Data

- Two covariates, Z_1 and $Z_2 \sim N(0,1)$, effect $\beta_1 = \beta_2 = 0.5$
- T_1^* unit exponential with mass 0.7 at ∞ if $Z_1 = Z_2 = 0$
- **L uniform on $[0,3]$ with mass 0.4 at 0**

Results for β_1 (5000 runs):

	N	bias	p-value	$\widehat{\text{var}(\beta_1)}$	$\overline{\text{var}(\beta_1)}$	$P(0.5 \in \text{CI})$	p-value
standard	103	0.022	$<$0.001	0.073	0.072	0.956	0.056
	258	0.007	$<$0.001	0.026	0.026	0.953	0.38
sandwich	103	0.022	$<$0.001	0.073	0.069	0.935	$<$0.001
	258	0.007	$<$0.001	0.026	0.025	0.943	0.025

Difference in standard error



Why Does Sandwich Perform Worse?

- Only individuals without event are reweighted
- Sandwich asymptotically unbiased, but less efficient
- Hence smaller as well as larger values occur

Remarks and conclusions

- No special software needed for competing risks analyses on cumulative scale
  r^* also used in log-rank type test of Gray (1988)
- No correction for reweighting needed in estimate of standard error
- The Fine & Gray model for competing risks with left truncated data or time-dependent covariates
  Fine & Gray model with categorical deterministic time dependent covariates: can be analysed if seen as truncated data (as in Kaplan-Meier for standard survival data)