An Efficient Alternative to the Cox Model for Small Time-to-Event Trials

Devan V. Mehrotra

Merck Research Laboratories, Blue Bell, PA

e-mail: *devan_mehrotra@merck.com*

FDA/Industry Statistics Workshop

Bethesda, MD

September 18, 2003

Outline

- Two group survival study
- Cox proportional hazards model
- Generalized logrank (GLR) statistic
    - continuous data, example 1
    - grouped data, example 2
- Link between the Cox model & GLR approaches
- Simulation study
- Concluding remarks

Two Group Survival Study

- Subjects randomized to treatment A or B.
- Time-to-event data (e.g., time to death, stroke, MI, etc.)
- Proportional hazards assumption

$h_j(t) = h_0(t)e^{\beta Z_j}$ = hazard for subject j at time t

$$Z_j = \begin{cases} 1 & \text{grp. A subject} \\ 0 & \text{grp. B subject} \end{cases}$$

At each t, $\dfrac{\text{hazard for grp. A}}{\text{hazard for grp. B}} = e^{\beta} = \theta$ (say)

θ = relative risk parameter, $\beta = \ln(\theta)$

- Inference : How to test $H_0\!:\beta = \beta_0$ (or $H_0\!:\theta = \theta_0$)?

Estimation : How to estimate β (or θ)?

Cox Model Approach

Notation

- $t_1 < t_2 < \ldots < t_k$ = event times

  d_i subjects fail at t_i (d_{iA} from A, d_{iB} from B)

  n_i subjects at risk just before t_i (n_{iA} in A, n_{iB} in B)

  $R(t_i)$ = set of subjects at risk at t_i

  $D(t_i)$ = set of subjects who have event at t_i

  $R(t_i, d_i)$ = set of d_i subjects chosen without replacement from $R(t_i)$

  $Z_{i1}, Z_{i2}, \ldots, Z_{id_i}$ = group indicators for the d_i failures at t_i

  $$S_i = \sum_{j=1}^{d_i} Z_{ij} = d_{iA}$$

Cox Model: No Ties

- Only one failure (event) at each t_i
- Cox Partial likelihood (1972)

$$L(\beta) = \prod_{i=1}^{k} L_i(\beta)$$

where

$$L_i(\beta) = \frac{e^{Z_i \beta}}{\sum_{j \in R(t_i)} e^{Z_j \beta}}$$

$$= \begin{cases} \dfrac{e^{\beta}}{n_{iA} e^{\beta} + n_{iB}} & \text{if event at } t_i \text{ is from group A} \\[2ex] \dfrac{1}{n_{iA} e^{\beta} + n_{iB}} & \text{if event at } t_i \text{ is from group B} \end{cases}$$

Cox Model: Ties

- d_i subjects fail "at t_i" (d_{iA} in A, d_{iB} in B) $d_i > 1$ for at least one i

  Various approximations of $L_i(\beta)$ are used to handle ties.

<u>Popular Tie-Handling Approximations</u>

- Cox (1972)

$$L_i^C(\beta) = \frac{\prod_{j=1}^{d_i} e^{Z_{ij}\beta}}{\sum_{l \in R(t_i, d_i)} e^{S_l \beta}} \qquad \text{(discrete partial likelihood)}$$

- Kalbfleisch & Prentice (1973)

$$L_i^{KP}(\beta) = \sum_{x=0}^{d_{iA}} \sum_{y=0}^{d_{iB}} (-1)^{x+y} \binom{d_{iA}}{x} \binom{d_{iB}}{y} \frac{e^{\beta}(n_{iA} - d_{iA}) + (n_{iB} - d_{iB})}{e^{\beta}(n_{iA} - d_{iA} + x) + (n_{iB} - d_{iB} + y)}$$

(simplification of the integral representation used by SAS PHREG)

Cox Model: Ties (cont'd)

- Breslow (1974)

$$L_i^B(\beta) = \frac{d_i! e^{S_i \beta}}{\left[\sum_{l \in R(t_i)} e^{Z_l \beta} \right]^{d_i}}$$

{<u>Default</u> in many software packages, including SAS}

- Efron (1977)

$$L_i^E(\beta) = \frac{d_i! e^{S_i \beta}}{\prod_{m=1}^{d_i} \left[\sum_{l \in R(t_i)} e^{S_l \beta} - \frac{(m-1)}{d_i} \sum_{l \in D(t_i)} e^{Z_l \beta} \right]}$$

{<u>Default</u> in S-Plus}

<u>No ties</u> $\Rightarrow$ $L_i^C(\beta) = L_i^{KP}(\beta) = L_i^B(\beta) = L_i^E(\beta) = L_i(\beta)$

Cox Model: Estimation

- Let $S(\beta) = \dfrac{\partial \log L(\beta)}{\partial \beta}$

  $I(\beta) = -\dfrac{\partial^2 \log L(\beta)}{\partial \beta^2}$

  $U(\beta) = \dfrac{\{S(\beta)\}^2}{I(\beta)}$

- Maximum partial likelihood estimate of β ($\tilde{\beta}_c$, say)
  - iterative solution of $S(\beta) = 0$
  - $V(\tilde{\beta}_c) = [I(\tilde{\beta}_c)]^{-1}$

Cox Model: Inference

- $H_0: \beta = \beta_0$ vs. $H_0: \beta \neq \beta_0$

Wald Test : $\dfrac{\left(\hat{\beta}_{CPH} - \beta_0\right)^2}{V\left(\hat{\beta}_{CPH}\right)}$

Score Test : $\dfrac{\left[S(\beta_0)\right]^2}{I(\beta_0)}$

Reference distribution for each test is $\chi^2_{(1)} = F(1, \infty)$

Cox Model: Inference (cont'd)

- $H_0 : \beta = 0$ vs. $H_0 : \beta \neq 0$

  (i.e. relative risk = 1 vs. ≠ 1)

  $$\text{“Logrank” statistic} = \frac{\left[\sum_{i=1}^{k} \left(d_{iA} - E_{iA}\right)\right]^2}{\sum_{i=1}^{k} V_{iA}} \quad \text{(Mantel, 1966)}$$

  $$E_{iA} = d_i \frac{n_{iA}}{n_i} = E\left(d_{iA} \middle| n_{iA}, n_{iB}, d_i, \beta = 0\right)$$

  $$V_{iA} = \frac{n_{iA} n_{iB} d_i \left(n_i - d_i\right)}{n_i^2 \left(n_i - 1\right)} = V\left(d_{iA} \middle| n_{iA}, n_{iB}, d_i, \beta = 0\right)$$

  Note: Mantel's LR statistic = <u>score statistic</u> based on Cox's discrete partial likelihood.

Generalized Logrank Statistic

Mehrotra and Roth (2001)

- Let $t_1 < t_2 < \ldots < t_k$ be the event times. At t_i:

	Fail	Survive	Total
Group A	d_{iA}	$n_{iA} - d_{iA}$	n_{iA}
Group B	d_{iB}	$n_{iB} - d_{iB}$	n_{iB}
Total	d_i	$n_i - d_i$	n_i

- If $d_{iB} \sim B(n_{iB}, p_i)$, under proportional hazards $d_{iA} \dot{\sim} B(n_{iA}, \theta p_i)$

$$P(D_{iA} = d_{iA} | d_i, n_{iA}, n_{iB}, p_i, \theta)$$

$$= \frac{\binom{n_{iA}}{d_{iA}}\binom{n_{iB}}{d_i - d_{iA}} \theta^{d_{iA}} (1-\theta p_i)^{n_{iA}-d_{iA}} (1-p_i)^{n_{iB}-d_i+d_{iA}}}{\sum_{j \in G_i} \binom{n_{iA}}{j}\binom{n_{iB}}{d_i - j} \theta^{j} (1-\theta p_i)^{n_{iA}-j} (1-p_i)^{n_{iB}-d_i+j}} \quad (1)$$

where $G_i = \{ j : \max(0, d_i - n_{iB}) \le j \le \min(d_i, n_{iA}) \}$

<u>Note</u>: If $\theta = 1$, then (1) does not involve p_i.

GLR: Inference

- Let $E_{iA}(n_{iA}, n_{iB}\theta, p_i)$ & $V_{iA}(n_{iA}, n_{iB}, \theta, p_i)$ = mean & variance of (1). For e.g., with no ties ($d_i = 1$):

$$E_{iA}(n_{iA}, n_{iB}\theta, p_i) = \frac{n_{iA}\theta(1-p_i)}{n_{iA}\theta(1-p_i) + n_{iB}(1-\theta p_i)}$$

$$V_{iA}(n_{iA}, n_{iB}, \theta, p_i) = \frac{n_{iA} n_{iB}\theta(1-p_i)(1-\theta p_i)}{\left[n_{iA}\theta(1-p_i) + n_{iB}(1-\theta p_i)\right]^2}.$$

- Proposed GLR statistic for $H_0: \theta = \theta_0$ vs. $H_A: \theta \neq \theta_0$

$$GLR(\theta_0, \underline{\tilde{p}}) = \frac{\left\{\sum_{i=1}^{k}\left[d_{iA} - E_{iA}(n_{iA}, n_{iB}, \theta_0, \tilde{p}_i)\right]\right\}^2}{\sum_{i=1}^{k} V_{iA}(n_{iA}, n_{iB}, \theta_0, \tilde{p}_i)}$$

where $\tilde{p}_i$ is an estimate of p_i & $\underline{\tilde{p}} = (\tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_k)$.

- $GLR(1, \underline{\tilde{p}})$ = Mantel's (1966) logrank statistic; it does not require estimation of $\underline{p}$.

GLR: Inference (cont'd)

- How to estimate p_i?

  Conditional Approach

  Find p_i that maximizes $\prod_i P\left(D_{iA} = d_{iA} \middle| d_i, n_{iA}, n_{iB}, p_i, \theta\right)$.

  For e.g., with no ties:

  $\tilde{p}_{i,\theta} = 0$ if $(d_{iA} = 1 \cap \theta < 1)$ or $(d_{iA} = 0 \cap \theta > 1)$

  $\tilde{p}_{i,\theta} = \min\left(1, \theta^{-1}\right)$ if $(d_{iA} = 1 \cap \theta > 1)$ or $(d_{iA} = 0 \cap \theta < 1)$.

  Note: d_i contains information about p_i that is 'wasted' if we condition on d_i.

GLR: Inference (cont'd)

Unconditional Approach

Numerator of (1) $\equiv L(p_i|\theta) \propto B(n_{iA}, \theta p_i) \times B(n_{iB}, p_i)$.

Find p_i that maximizes $\prod_i L(p_i|\theta)$. We get:

$$\tilde{p}_i \equiv \tilde{p}_{i,\theta} = \frac{x_i - \sqrt{x_i^2 - 4(n_{iA} + n_{iB})d_i\theta}}{2(n_{iA} + n_{iB})\theta}$$

where $x_i = \theta(n_{iA} + d_i - d_{iA}) + n_{iB} + d_{iA}$.

GLR: Inference (cont'd)

- Let $\underline{\tilde{p}}(\theta) = \left(\tilde{p}_{1,\theta}, \tilde{p}_{2,\theta}, \ldots, \tilde{p}_{k,\theta}\right)$, unconditional approach.
- Reference distribution for $GLR\left[\theta_0, \underline{\tilde{p}}(\theta_0)\right] = ?$

  Proposed approximation:

  $GLR\left[\theta_0, \underline{\tilde{p}}(\theta_0)\right] \sim_{null} F\left(1, k^*\right)$

  where $k^* = \sum_{i=1}^{k} \min\left(d_i, n_i - d_i, n_{iA}, n_{iB}\right)$.

Remarks

- With no ties, k^* = # of 'informative' tables.
- $F\left(1, k^*\right) \to \chi_1^2$ as $k^* \to \infty$.

GLR: Estimation

- Small values of $GLR\left[\theta_0, \underline{\tilde{p}}(\theta_0)\right]$ support $H_0: \theta = \theta_0$.

  The GLR-based estimator of θ, $\tilde{\theta}_{GLR}$, is the θ which satisfies:

$$GLR\left[\tilde{\theta}_{GLR}, \underline{\tilde{p}}\left(\tilde{\theta}_{GLR}\right)\right] = \inf_{\theta} GLR\left[\theta, \underline{\tilde{p}}(\theta)\right].$$

- $100(1-\alpha)\%$ confidence interval for θ

$$\theta_{GLR}^{L} = \inf_{\theta}\left\{\theta : GLR\left[\theta, \underline{\tilde{p}}(\theta)\right] \le F_{\alpha}\left(1, k^{*}\right)\right\}$$

$$\theta_{GLR}^{U} = \sup_{\theta}\left\{\theta : GLR\left[\theta, \underline{\tilde{p}}(\theta)\right] \le F_{\alpha}\left(1, k^{*}\right)\right\}$$

Example 1

Survival times (days) of 30 patients with cervical cancer

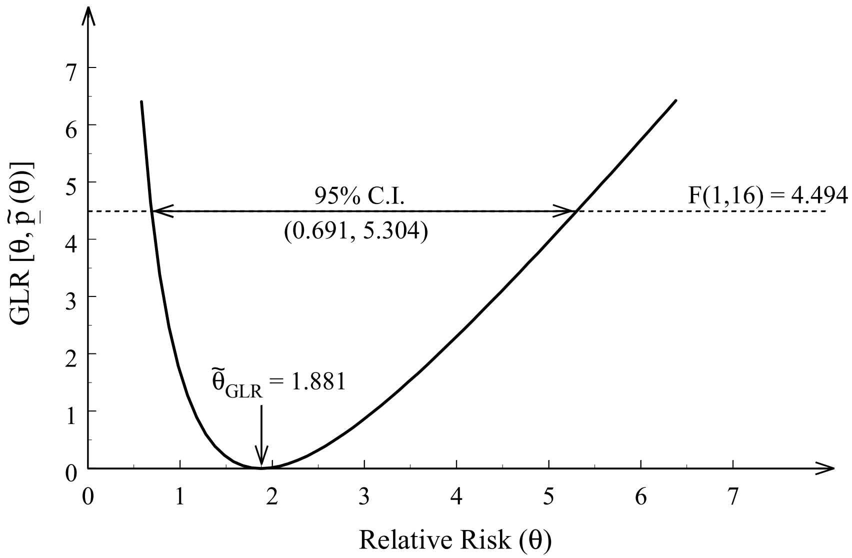
Group A (control therapy, $N_A = 16$)

90, 142, 150, 269, 291, 468+, 680, 837, 890+, 1037, 1090+, 1113+, 1153, 1297, 1429, 1577+

Group B (new therapy, $N_B = 14$)

272, 362, 373, 383+, 519+, 563+, 650+, 827, 919+, 978+, 1100+, 1307, 1360+, 1476+

Source: Pg. 69, Parmar and Machin (1995); + denotes censored observation



GLR method: $\tilde{\theta}_{GLR} = 1.88$, 95% C.I. = (0.69, 5.30)

Cox Model: $\tilde{\theta}_C = 2.00$, 95% C.I. = (0.69, 5.80)

GLR: Grouped Data

- GLR statistic will not work well when many ties result from a coarse grouping of continuous data. (e.g., rounding to weeks, months, etc.)

- Two simple extensions of our GLR statistic for grouped data: GLR^{KP} and GLR^{E}

  In each case, replace E_{iA} & V_{iA} in the GLR statistic (on pg. 12) with $\overline{E}_{iA}$ & $\overline{V}_{iA}$, respectively.

  $\overline{E}_{iA}$ = mean of E_{iA} over all possible orderings of events between t_{i-1} and t_i.

  $\overline{V}_{iA}$ = mean of V_{iA} over all possible orderings of events between t_{i-1} and t_i.

- See Appendix for details.

Example 2

Time (in weeks) to a particular adverse event

Group A (old therapy, N_A=20)
2, **2**, 4+, **8**, 8+, **12**, **12**, **12**, **12**, 12+, 12+, 16+, 16+, 20+, 24+, 24+, 28+, 28+, **36**, **36**

Group B (new therapy, N_B=20)
4+, 4+, 4+, 4+, **8**, 12+, 12+, 16+, 16+, 16+, 16+, 20+, 20+, 24+, 28+, 28+, 32+, 32+, **36**, 36+

Note: scheduled post-randomization visits at wks 2, 4 and then every 4 wks; + censored observation

Estimated Relative Risks (95% C.I.s)

Cox Model-Based Approaches			GLR Approaches		
Breslow	4.45	(0.96, 20.60)			
Cox	8.71	(1.06, 71.60)			
Efron	5.12	(1.10, 23.86)	3.76	(1.03, 18.01)	GLR^E
K & P	7.95	(1.04, 60.85)	4.71	(1.05, 25.01)	GLR^{KP}

- To determine sample size for a "confirmatory study", what is the best guess of the true relative risk? Note: sample size $\propto \left(\frac{1+RR}{1-RR}\right)^2$

Link Between GLR and Cox Model Approaches

- If $p_i = 0 \ \forall i$, GLR = <u>score statistic</u> based on the Cox discrete partial likelihood.

  If $p_{i,j} = 0 \ \forall i, j$, GLR^{E} = <u>score statistic</u> based on the Efron partial likelihood.

  If $p_{i,x+y} = 0 \ \forall i, x, y$, GLR^{KP} ≈ <u>score statistic</u> based on the Kalbfleisch & Prentice partial likelihood.

- **When all the risk sets (i.e. n_i's) are very large**

  GLR ≈ Cox

  GLR^{KP} ≈ K & P

  GLR^{E} ≈ Efron

  (since $\underline{\tilde{p}} \approx \underline{0}$ with large risk sets)

- This suggests that GLR and Cox model will be asymptotically similar.

Simulation Study

- N subjects per treatment (total size = 2N subjects).
- E = entry time into trial ~ Uniform(0,1). Analysis at *T*.

  S = true survival time ~ $Weibull(scale,\ shape)$

  Trt A: $\left(scale = \sqrt{0.5e^{\beta}}, shape = 2\right)$ $\quad \left\{h(t) = te^{\beta}\right\}$

  Trt B: $\left(scale = \sqrt{0.5},\ shape = 2\right)$ $\quad \left\{h(t) = t\right\}$

  Y = observed survival time = $\min(S, T - E)$.

- For a given *T*, let $\pi_j(T)$ = probability that a subject in group *j* was censored. For N_A=N_B, the mean probability of censoring was $\pi(T) = 0.5\{\pi_A(T) + \pi_B(T)\}$. We studied three values of *T*, namely, T_1, T_2 and T_3, such that $\pi(T_1) = 0\%$, $\pi(T_2) = 25\%$ and $\pi(T_3) = 50\%$.
- Continuous data (no ties), Grouped data (ties, data rounded to nearest 0.10)

Continuous Data (No ties)

BIAS

		% censoring					
		0 %		25 %		50 %	
2N	β	Cox	GLR	Cox	GLR	Cox	GLR
40	0.0	.000	.000	.000	.000	.004	.004
	0.6	.025	-.023	.022	-.022	.024	-.020
	1.6	.075	-.048	.071	-.056	.082	-.067
80	0.0	.006	.005	.004	.004	.007	.007
	0.6	.016	-.011	.013	-.010	.009	-.014
	1.6	.032	-.029	.028	-.036	.031	-.046
200	0.0	.003	.003	.001	.001	.002	.002
	0.6	.007	-.005	.003	-.007	.007	-.002
	1.6	.013	-.012	.012	-.015	.012	-.021

2000 replications

Range of |%BIAS| for $\beta > 0$

Cox	0.5% to 5.1%
GLR	0.4% to 4.2%

Continuous Data (no ties)

Mean Squared Error (MSE) & Relative Efficiency

		% censoring								
		0 %			25 %			50 %		
2N	β	Cox	GLR	%RE	Cox	GLR	%RE	Cox	GLR	%RE
40	0.0	.121	.102	**119**	.143	.125	**115**	.211	.184	**114**
	0.6	.126	.109	**115**	.152	.132	**116**	.221	.190	**117**
	1.6	.196	.160	**123**	.211	.172	**123**	.328	.261	**126**
80	0.0	.055	.050	**111**	.070	.065	**108**	.109	.101	**107**
	0.6	.059	.055	**108**	.074	.068	**108**	.112	.103	**108**
	1.6	.089	.082	**109**	.097	.089	**109**	.150	.133	**113**
200	0.0	.021	.020	**105**	.028	.027	**103**	.039	.038	**103**
	0.6	.023	.023	**103**	.029	.028	**103**	.042	.040	**103**
	1.6	.033	.032	**103**	.036	.035	**103**	.054	.051	**105**

%RE = 100x(MSE_{Cox})/(MSE_{GLR}); 2000 replications

Continuous Data (no ties)

Coverage of 95% C.I. & Relative Width

		% censoring								
		0 %			25 %			50 %		
2N	β	Cox	GLR	R.W.	Cox	GLR	R.W.	Cox	GLR	R.W.
40	0.0	94.7	95.2	.92	94.9	95.3	.95	95.5	96.0	.97
	0.6	95.1	95.1	.92	95.3	95.7	.95	96.5	96.7	.96
	1.6	96.5	94.6	.90	96.3	94.7	.91	97.6	95.8	.93
80	0.0	94.3	94.4	.95	95.2	95.3	.97	95.4	95.8	.98
	0.6	94.8	94.2	.96	95.0	95.1	.97	95.0	95.1	.98
	1.6	95.0	94.3	.95	95.2	93.9	.95	96.2	94.4	.95
200	0.0	94.3	94.3	.98	94.7	94.7	.99	95.8	95.9	.99
	0.6	94.3	94.1	.98	95.3	95.0	.99	95.4	95.4	.99
	1.6	94.9	94.2	.98	95.1	94.2	.98	95.5	94.9	.98

R.W. = $(\text{width}_{GLR})/(\text{width}_{Cox})$; 2000 replications

Concluding Remarks

- Continuous data → GLR approach can provide substantial gains in efficiency relative to the Cox model in small to moderate sized trials (< 200 subjects).

- Grouped data → GLR^{KP} and GLR^{E} are also notably more efficient than their Cox model counterparts.

- GLR and Cox model approaches are asymptotically similar, for both continuous and grouped data.

- Research has been extended to include stratification (Mehrotra and Zou, manuscript in preparation).

APPENDIX

GLR: Grouped Data (Development of Formulas)

<u>Example</u>

3 events between t_{i-1} and t_i $\left(d_{iA}=2,\ d_{iB}=1\right)$

True event times: $t_{i-1} < t_{i,1} < t_{i,2} < t_{i,3} < t_i$

Possible oderings	# at risk in Group (A, B) just before $t_{i,j}$ j = 1	j = 2	j = 3
A→A→B	$\left(n_{iA}, n_{iB}\right)$	$\left(n_{iA}-1, n_{iB}\right)$	$\left(n_{iA}-2, n_{iB}\right)$
A→B→A	$\left(n_{iA}, n_{iB}\right)$	$\left(n_{iA}-1, n_{iB}\right)$	$\left(n_{iA}-1, n_{iB}-1\right)$
B→A→A	$\left(n_{iA}, n_{iB}\right)$	$\left(n_{iA}, n_{iB}-1\right)$	$\left(n_{iA}-1, n_{iB}-1\right)$
Average # at risk	$\left(n_{iA}, n_{iB}\right)$	$\left(n_{iA}-\frac{2}{3}, n_{iB}-\frac{1}{3}\right)$	$\left(n_{iA}-\frac{4}{3}, n_{iB}-\frac{2}{3}\right)$
Group B Binomial Probability	$p_{i,1}$	$p_{i,2}$	$p_{i,3}$

<u>For GLR^{KP}</u>

$$
\begin{aligned}
\bar{E}_{iA}^{KP} = \quad & \frac{1}{3}\left[3E_{iA}\left(n_{iA}, n_{iB}, \theta, p_{i,1}\right)\right] \\
+ \quad & \frac{1}{3}\left[2E_{iA}\left(n_{iA}-1, n_{iB}, \theta, p_{i,2}\right) + E_{iA}\left(n_{iA}, n_{iB}-1, \theta, p_{i,2}\right)\right] \\
+ \quad & \frac{1}{3}\left[2E_{iA}\left(n_{iA}-1, n_{iB}-1, \theta, p_{i,3}\right) + E_{iA}\left(n_{iA}-2, n_{iB}, \theta, p_{i,3}\right)\right]
\end{aligned}
$$

GLR: Grouped Data (cont'd)

<u>Example (cont'd)</u>

<u>For GLR^{KP}</u>

$$\begin{aligned}\bar{V}_{iA}^{KP} = \; & \frac{1}{3}\left[3V_{iA}\left(n_{iA}, n_{iB}, \theta, p_{i,1}\right)\right] \\ + \; & \frac{1}{3}\left[2V_{iA}\left(n_{iA}-1, n_{iB}, \theta, p_{i,2}\right) + V_{iA}\left(n_{iA}, n_{iB}-1, \theta, p_{i,2}\right)\right] \\ + \; & \frac{1}{3}\left[2V_{iA}\left(n_{iA}-1, n_{iB}-1, \theta, p_{i,3}\right) + V_{iA}\left(n_{iA}-2, n_{iB}, \theta, p_{i,3}\right)\right]\end{aligned}$$

<u>For GLR^{E}</u>

$$\begin{aligned}\bar{E}_{iA}^{E} = \; & E_{iA}\left(n_{iA}, n_{iB}, \theta, p_{i,1}\right) \\ + \; & E_{iA}\left(n_{iA}-\frac{2}{3}, n_{iB}-\frac{1}{3}, \theta, p_{i,2}\right) \\ + \; & E_{iA}\left(n_{iA}-\frac{4}{3}, n_{iB}-\frac{2}{3}, \theta, p_{i,3}\right)\end{aligned}$$

$$\begin{aligned}\bar{V}_{iA}^{E} = \; & V_{iA}\left(n_{iA}, n_{iB}, \theta, p_{i,1}\right) \\ + \; & V_{iA}\left(n_{iA}-\frac{2}{3}, n_{iB}-\frac{1}{3}, \theta, p_{i,2}\right) \\ + \; & V_{iA}\left(n_{iA}-\frac{4}{3}, n_{iB}-\frac{2}{3}, \theta, p_{i,3}\right)\end{aligned}$$

GLR: Grouped Data (cont'd)

General Formulas

For GLR^{KP}

$$\overline{E}_{iA}^{KP} = \frac{\sum_{x=0}^{d_{iA}} \sum_{y=0}^{d_{iB}} I_{x+y} \binom{x+y}{x} \binom{d_i - x - y}{d_{iA} - x} E_{iA}\left(n_{iA} - x, n_{iB} - y, \theta, p_{i,x+y}\right)}{\binom{d_i}{d_{iA}}}$$

$$\overline{V}_{iA}^{KP} = \frac{\sum_{x=0}^{d_{iA}} \sum_{y=0}^{d_{iB}} I_{x+y} \binom{x+y}{x} \binom{d_i - x - y}{d_{iA} - x} V_{iA}\left(n_{iA} - x, n_{iB} - y, \theta, p_{i,x+y}\right)}{\binom{d_i}{d_{iA}}}$$

where

$p_{i,x+y}$ = binomial failure prob. for grp. B at $t_{i,j}$, with $j = x + y + 1$

$I_{x+y} = 0$ if $x + y = d_i$; 1 otherwise.

For GLR^{E}

$$\overline{E}_{iA}^{E} = \sum_{j=1}^{d_i} E_{iA}\left(n_{iA} - (j-1)\frac{d_{iA}}{d_i}, n_{iB} - (j-1)\frac{d_{iB}}{d_i}, \theta, p_{i,j} \right)$$

$$\overline{V}_{iA}^{E} = \sum_{j=1}^{d_i} V_{iA}\left(n_{iA} - (j-1)\frac{d_{iA}}{d_i}, n_{iB} - (j-1)\frac{d_{iB}}{d_i}, \theta, p_{i,j} \right)$$

GLR: Grouped Data (cont'd)

How to estimate the $p_{i,j}$'s?

"Average 2x2 Table" at $t_{i,j}$

	Fail	Survive	Total
Group A	$\frac{d_{iA}}{d_i}$	$n_{iA} - j\frac{d_{iA}}{d_i}$	$n_{iA} - (j-1)\frac{d_{iA}}{d_i}$
Group B	$\frac{d_{iB}}{d_i}$	$n_{iB} - j\frac{d_{iB}}{d_i}$	$n_{iB} - (j-1)\frac{d_{iB}}{d_i}$
Total	1	$n_i - j$	$n_i - (j-1)$

Analogous to GLR approach for continuous data:

$$\tilde{p}_{i,j} \equiv \tilde{p}_{i,j,\theta} = \frac{h_i - \sqrt{h_i^2 - 4(n_{iA} + n_{iB} - j + 1)\theta}}{2(n_{iA} + n_{iB} - j + 1)\theta}$$

where

$$h_i = \theta\left[n_{iA} + \frac{d_{iB}}{d_i} - (j-1)\frac{d_{iA}}{d_i}\right] + n_{iB} + \frac{d_{iA}}{d_i} - (j-1)\frac{d_{iB}}{d_i}$$

Grouped Data (ties) – Simulation Results

Cox Model: Comparing Popular Tie-Handling Approximations

- N=100/grp, $\beta = 1.2$, no censoring

%BIAS (for β)

Interval Width	Cox	Breslow	Efron	K & P
0.00	0.6	0.6	0.6	0.6
0.02	2.3	-0.9	0.6	0.7
0.10	8.7	-6.9	0.5	1.0
0.25	19.3	-17.8	-1.8	0.7

2000 replications

- With many ties …

  Computing : (Breslow, Efron) << (K & P, Cox)

  Bias : (K & P, Efron) << (Cox, Breslow)

Grouped Data (ties)

BIAS

		0 %		25 %		50 %	
2N	β	Efron*	GLR^E	Efron*	GLR^E	Efron*	GLR^E
40	0.0	-.005	-.004	.006	.006	.015	.014
	0.6	.028	-.024	.012	-.031	.022	-.019
	1.6	.037	-.087	.038	-.087	.050	-.094
80	0.0	.006	.006	-.002	-.002	-.005	-.005
	0.6	.010	-.018	.010	-.012	.104	-.008
	1.6	.020	-.042	.019	-.044	.023	-.050
200	0.0	-.004	-.004	-.003	-.003	-.003	-.003
	0.6	.007	-.005	-.002	-.011	-.003	-.011
	1.6	.007	-.018	.008	-.018	.014	-.016

* Efron's (1977) partial likelihood; 2000 replications

Range of |%BIAS| for $\beta > 0$

Efron	0.1% to 4.6%	GLR^E	0.8% to 5.9%
K & P	0.1% to 6.7%	GLR^{KP}	0.7% to 5.2%

Grouped Data (ties)

Mean Squared Error (MSE) & Relative Efficiency

		% censoring								
		0 %			25 %			50 %		
2N	β	Efron*	GLR^E	% RE	Efron*	GLR^E	% RE	Efron*	GLR^E	% RE
40	0.0	.119	.099	**120**	.141	.123	**115**	.215	.190	**113**
	0.6	.124	.108	**115**	.147	.128	**115**	.240	.208	**115**
	1.6	.190	.160	**118**	.215	.180	**119**	.329	.270	**122**
80	0.0	.054	.048	**112**	.066	.062	**107**	.105	.099	**106**
	0.6	.059	.056	**107**	.071	.066	**107**	.103	.095	**108**
	1.6	.089	.083	**107**	.091	.085	**107**	.153	.137	**111**
200	0.0	.021	.020	**106**	.028	.027	**103**	.041	.040	**103**
	0.6	.023	.022	**103**	.027	.026	**103**	.041	.040	**103**
	1.6	.033	.032	**103**	.037	.036	**103**	.056	.053	**105**

* Efron's (1977) partial likelihood; %RE = 100x(MSE_{Efron})/($MSE_{GLR}{}^{E}$); 2000replications

- MSEs → (GLR^E, GLR^{KP}) < (Efron, K & P)

Grouped Data (ties)

Coverage of 95% C.I. & Relative Width

		% censoring								
		0 %			25 %			50 %		
2N	β	Efron*	GLR^E	R.W.	Efron*	GLR^E	R.W.	Efron*	GLR^E	R.W.
40	0.0	94.5	95.0	.91	95.6	96.1	.95	96.5	96.9	.97
	0.6	95.1	94.7	.92	95.5	95.5	.95	94.6	95.1	.96
	1.6	94.9	(92.4)	.89	95.2	94.2	.91	96.1	94.6	.93
80	0.0	95.3	95.4	.95	95.8	96.0	.98	94.9	95.4	.98
	0.6	94.4	94.2	.95	95.1	95.1	.97	96.1	96.5	.98
	1.6	94.9	93.7	.95	95.7	95.0	.95	95.2	94.0	.95
200	0.0	95.1	95.2	.97	94.1	94.2	.99	94.8	95.0	.99
	0.6	95.1	94.5	.98	95.6	95.4	.99	95.2	95.2	.99
	1.6	94.8	94.3	.98	94.8	94.6	.98	94.9	94.8	.98

* Efron's (1977) partial likelihood; R.W. = $(\text{width}_{GLR}{}^E)/(\text{width}_{Efron})$; 2000 replications

- GLR^{KP} provides slightly better coverage than GLR^E

REFERENCES

Breslow, N. E. (1974). Covariance analysis of censored survival data. Biometrics, 30, 89-99.

Cox, D. R. (1972). Regression models and life-tables (with discussion). Journal of the Royal Statistical Society, B, 34, 187-220.

Efron, B. (1977). Efficiency of Cox's likelihood function for censored data. Journal of the American Statistical Association, 72, 557-565.

Kalbfleisch, J. D. and Prentice, R. L. (1973). Marginal likelihood based on Cox's regression and life model. Biometrika, 60, 267-278.

Mantel, N. (1966). Evaluation of survival data and two new rank order statistics arising in its consideration. Cancer Chemotherapy Reports, **50**, 163-170.

Mehrotra, D. V. and Roth, A. J. (2001). Relative risk estimation and inference using a generalized logrank statistic. Statistics in Medicine, 20, 2099-2113.

Parmar, M. K. B. and Machin, D. Survival Analysis: A Practical Approach, John Wiley and Sons: Chichester (1995).

Peto, R. and Peto, J. (1972). Asymptotically efficient rank invariant test procedures (with discussion). Journal of the Royal Statistical Society, Series A, **135**, 185-206.