Statistical Issues in the Use of Composite Endpoints in Clinical Trials

Longyang Wu and Richard Cook

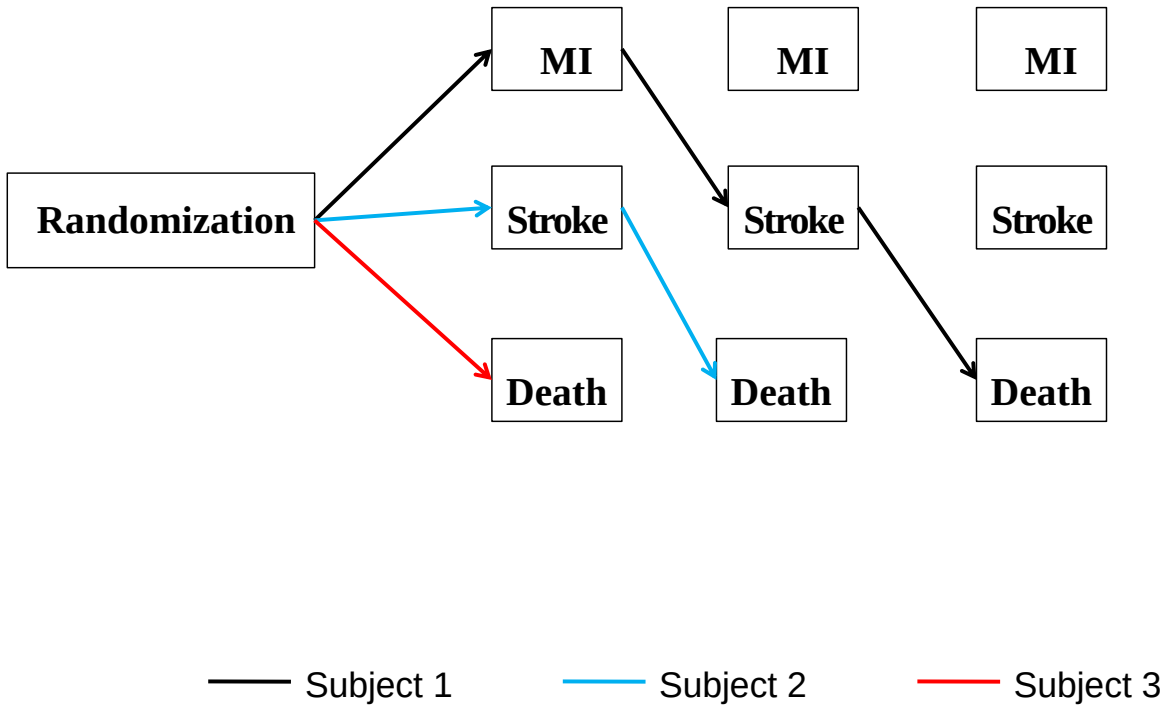
Statistics and Actuarial Science, University of Waterloo

May 14, 2010

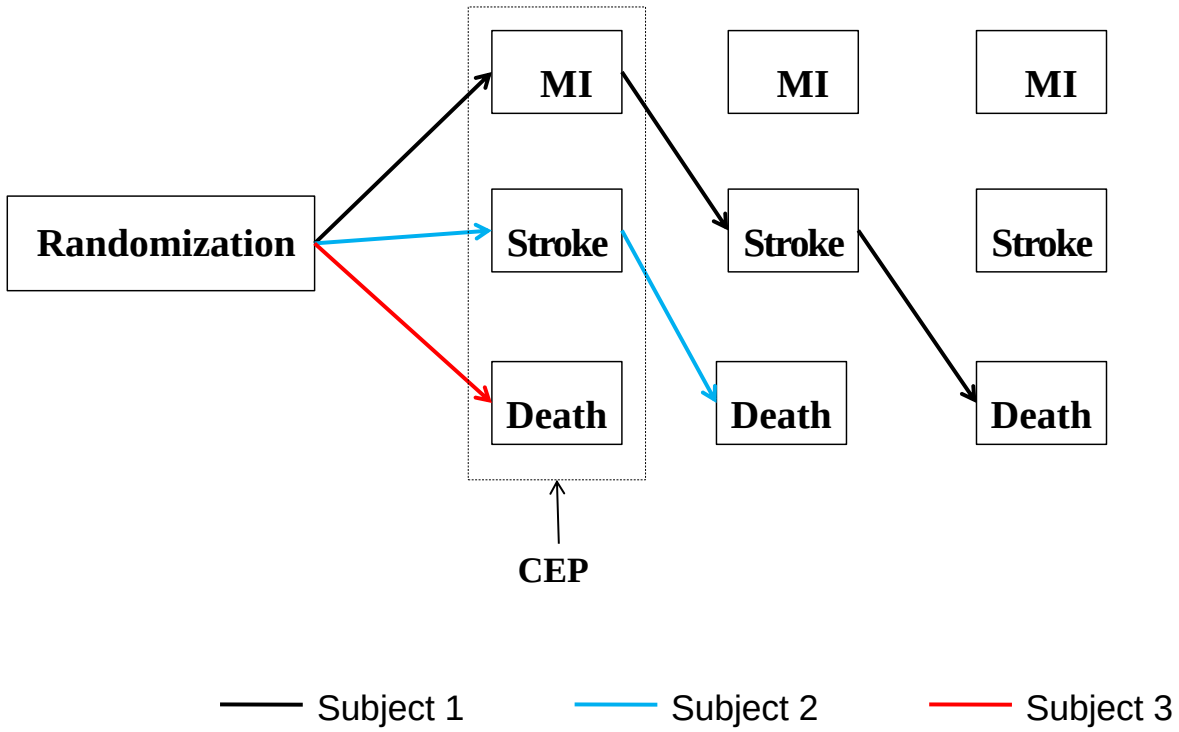
COMPOSITE ENDPOINTS

- Composite endpoints (CEP) in clinical trials:
  - CEP is defined as a group of endpoints that together form a single endpoint in a trial (Cannon, 1997).
  - Those individual endpoints are referred to as components of the CEP.
- The HOPE study (Yusuf *et al.*, 2000) investigated the effects of ramipril on cardiovascular events in high-risk patients with left ventricular dysfunction.
- CEP in HOPE:
  - myocardial infarction (MI)
  - stroke
  - cardiovascular death
- Patients who have any of these three events during the follow-up are considered to have experienced the CEP.

An Example of CEP



An Example of CEP



Our Present Objectives

- The focus of this study will be on CEP of time event data.
- We used models, asymptotic theory and empirical studies formally investigate the behaviour of estimation of treatment effects.
- Use simulations to study the implications of using CEP in the statistical power and sample size requirement.
- Investigate the validity of some of recommendations of CEP analysis.
- Study the effectiveness of alternative design and analysis

- Let T_i be the time to the CEP with $N_i(t) = I(T_i \leqslant t)$ and process history $\mathcal{H}_i(t) = \{N_i(u), Z_i, 0 < u < t\}$.
- Let $\Delta N_i(t) = N_i(t + \Delta t^-) - N_i(t^-)$ and $dN_i(t) = \lim_{\Delta t \downarrow 0} \Delta N_i(t)$.
- The hazard function for T_i is defined as

$$\lambda_i(t|\mathcal{H}_i(t)) = \mathscr{E}(dN_i(t)|\mathcal{H}_i(t)) = \lim_{\Delta t \downarrow 0} \frac{P(\Delta N_i(t) = 1|\mathcal{H}_i(t))}{\Delta t}$$

where $\mathscr{E}(\cdot)$ denotes the expectation operator with respect to the true process history.

Counting Process Notation: Review of Univariate Failure Times

- Let C_i denote censoring time and $X_i = \min(T_i, C_i)$.
- Let $Y_i(t) = I(X_i \geq t)$ be the at risk function.
- The observed counting process of T_i be $\bar{N}_i(t) = \int_0^t Y_i(u) dN_i(u)$
- Let $H_i(t) = \{N_i(u), Y_i(u), Z, 0 < u < t\}$ be the full history.
- Let $d\bar{N}_i(t) = \lim_{\Delta t \downarrow 0} \Delta \bar{N}_i(t)$.
- The hazard of the observed counting process of T_i is

$$\mathscr{E}(d\bar{N}_i(t) = 1 | H_i(t)) = Y_i(t)\lambda(t|Z_i)dt.$$

- The Cox model has the form

$$\lambda(t|Z_i) = \lambda_0(t)e^{\beta Z_i}$$

where $\lambda_0(t)$ is the baseline hazard function.

- The maximum partial likelihood estimation of β is obtained by solving:

$$U(\beta) = \sum_{i=1}^{n} \int_0^{\infty} \left(Z_i - \frac{S^{(1)}(\beta, t)}{S^{(0)}(\beta, t)} \right) d\bar{N}_i(t),$$

where $S^{(k)}(\beta, t) = \sum_{i=1}^{n} Y_i(t) Z_i^k \exp\{\beta Z_i\}, k = 0, 1.$

Composite Endpoint: Independent Case

- Let T_{i1} and T_{i2} be two independent failure times of subject i.
- T_{ik} has hazard functions $\lambda_k(t|Z_i)$, $k = 1, 2$.
- CEP analysis is based on $T_i = \min(T_{i1}, T_{i2})$ and $N_i(s) = I(s \leqslant T_i)$.
- The hazard function of CEP is

$$\mathscr{E}(dN_i(s)|H_i(s)) = \lambda_1(t|Z_i) + \lambda_2(t|Z_i).$$

- If $\lambda_{ik}(t|Z_i) = \lambda_{0k}(t)e^{\beta_k Z_i}$, $k = 1, 2$, then CEP has the hazard ratio

$$\frac{\lambda_{01}(t)e^{\beta_1} + \lambda_{02}(t)e^{\beta_2}}{\lambda_{01}(t) + \lambda_{02}(t)}.$$

Composite Endpoint: Independent Case

- Let T_{i1} and T_{i2} be two independent failure times of subject i.
- T_{ik} has hazard functions $\lambda_k(t|Z_i)$, $k = 1, 2$.
- CEP analysis is based on $T_i = \min(T_{i1}, T_{i2})$ and $N_i(s) = I(s \leqslant T_i)$.
- The hazard function of CEP is

$$\mathscr{E}(dN_i(s)|H_i(s)) = \lambda_1(t|Z_i) + \lambda_2(t|Z_i).$$

- If $\lambda_{ik}(t|Z_i) = \lambda_{0k}(t)e^{\beta_k Z_i}$, $k = 1, 2$, then CEP has the hazard ratio

$$\frac{\lambda_{01}(t)e^{\beta_1} + \lambda_{02}(t)e^{\beta_2}}{\lambda_{01}(t) + \lambda_{02}(t)}.$$

Remarks:

1.The proportional hazard assumption holds for CEP, if

(A.1) $\beta_1 = \beta_2$: the same treatment effect across components, or

(A.2) $\lambda_{01}(t) = \lambda_{02}(t)$: the same frequency of occurrence.

2. Otherwise, PH does not hold for CEP analysis.

Composite Endpoint: Independent Case

- Let T_{i1} and T_{i2} be two independent failure times of subject i.
- T_{ik} has hazard functions $\lambda_k(t|Z_i)$, $k = 1, 2$.
- CEP analysis is based on $T_i = \min(T_{i1}, T_{i2})$ and $N_i(s) = I(s \leqslant T_i)$.
- The hazard function of CEP is

$$\mathscr{E}(dN_i(s)|H_i(s)) = \lambda_1(t|Z_i) + \lambda_2(t|Z_i).$$

- If $\lambda_{ik}(t|Z_i) = \lambda_{0k}(t)e^{\beta_k Z_i}$, $k = 1, 2$, then CEP has the hazard ratio

$$\frac{\lambda_{01}(t)e^{\beta_1} + \lambda_{02}(t)e^{\beta_2}}{\lambda_{01}(t) + \lambda_{02}(t)}.$$

Remarks:

1.The proportional hazard assumption holds for CEP, if

(A.1) $\beta_1 = \beta_2$: the same treatment effect across components, or

(A.2) $\lambda_{01}(t) = \lambda_{02}(t)$: the same frequency of occurrence.

2. Otherwise, PH does not hold for CEP analysis.

What are we estimating in this case?

Composite Endpoint: Events Are Associated

- Typically components of CEP are not independent.
- Consider a joint survivor function of T_{i1} and T_{i2} via a Clayton copula:

$$(\mathcal{F}(t_{i1}|Z_i;\beta_1)^{-\phi} + \mathcal{F}(t_{i2}|Z_i;\beta_2)^{-\phi} - 1)^{-1/\phi}.$$

  - $\mathcal{F}(\cdot\,|\cdot\,)$ is the survivor function.
  - ϕ is the association parameter.
  - Kendall's $\tau = \frac{\phi}{2+\phi}$ is a common measure of association for T_{i1} and T_{i2}.
  - $\Theta = (\beta_1, \beta_2, \phi)$.

- Assume $T_k \sim \text{Exp}(\lambda_k e^{\beta_k Z})$, $k = 1, 2$
- Using the Clayton copula, the joint survivor function of T_1 and T_2 is

$$[\exp(\lambda_1 e^{\beta_1 Z} \phi t_1) + \exp(\lambda_2 e^{\beta_2 Z} \phi t_2) - 1]^{-1/\phi}.$$

- If $T = \min(T_1, T_2)$, $P(T \geqslant t|Z; \Theta) = \mathcal{F}(t|Z; \Theta)$ is

$$[\exp(\lambda_1 e^{\beta_1 Z} \phi t) + \exp(\lambda_2 e^{\beta_2 Z} \phi t) - 1]^{-1/\phi}$$

with density

$$f(t|Z; \Theta) = -\frac{\partial \mathcal{F}(t|Z; \Theta)}{\partial t}.$$

The hazard function of T is

$$\lambda_T(t|Z;\Theta) = \frac{f(t|Z;\Theta)}{\mathcal{F}(t|Z;\Theta)}.$$

The hazard ratio is :

$$\frac{f(t|Z=1;\Theta)}{f(t|Z=0;\Theta)} \times \frac{\mathcal{F}(t|Z=0;\Theta)}{\mathcal{F}(t|Z=1;\Theta)},$$

which is NOT a constant even if (A.1) and (A.2) are true.

The hazard function of T is

$$\lambda_T(t|Z;\Theta) = \frac{f(t|Z;\Theta)}{\mathcal{F}(t|Z;\Theta)}.$$

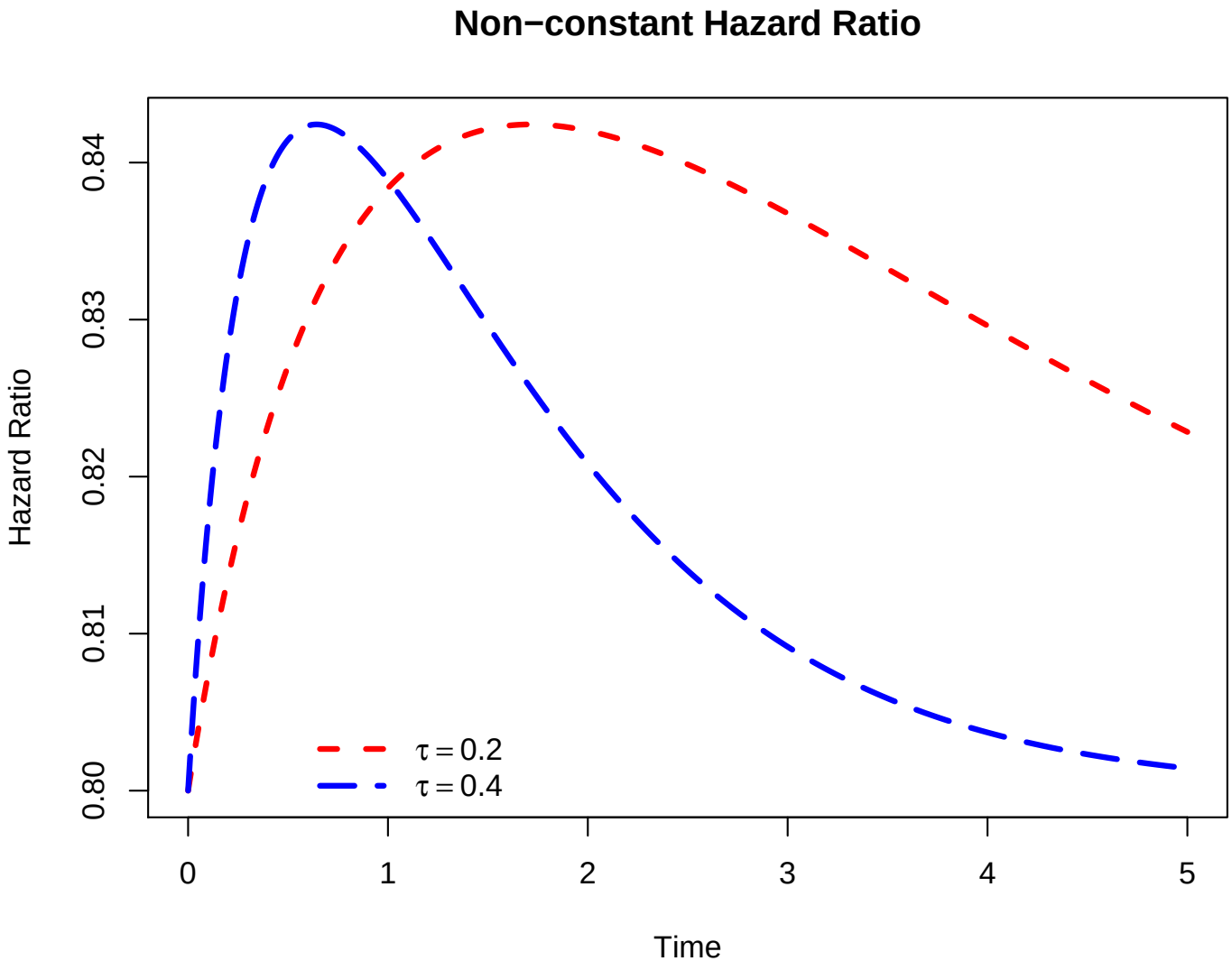
The hazard ratio is :

$$\frac{f(t|Z=1;\Theta)}{f(t|Z=0;\Theta)} \times \frac{\mathcal{F}(t|Z=0;\Theta)}{\mathcal{F}(t|Z=1;\Theta)},$$

which is NOT a constant even if (A.1) and (A.2) are true.

Remark:
In general, proportional hazards assumption does not hold for CEP, even if this assumption holds for individual components!

Non-constant Hazard Ratio: Composite Endpoint



- Treatment effects: $\beta_1 = \beta_2 = \log(0.8)$ and equal frequency of occurrence: $\lambda_1 = \lambda_2 = 1$.

Misspecified Cox Model: The Limiting Value

In CEP analysis, the solution of β to the partial likelihood score function

$$U(\beta) = \sum_{i=1}^{n} \int_0^{\infty} \left(Z_i - \frac{S^{(1)}(\beta, t)}{S^{(0)}(\beta, t)} \right) d\bar{N}_i(t)$$

is a consistent estimate of β^*, the solution to limiting score function

$$U(\beta^*) = \int_0^{\infty} \left(\mathscr{E}(\sum_{i=1}^{n} Z_i d\bar{N}_i(t)) - \frac{\mathscr{E}(S^{(1)}(\beta, t))}{\mathscr{E}(S^{(0)}(\beta, t))} \mathscr{E}(\sum_{i=1}^{n} d\bar{N}_i(t)) \right),$$

where the expectation $\mathscr{E}$ is with respect to the true model (e.g. White 1982, Struthers and Kalbfleisch, 1986).

Misspecified Cox Model: The Limiting Value

In CEP analysis, the solution of β to the partial likelihood score function

$$U(\beta) = \sum_{i=1}^{n} \int_0^{\infty} \left(Z_i - \frac{S^{(1)}(\beta, t)}{S^{(0)}(\beta, t)} \right) d\bar{N}_i(t)$$

is a consistent estimate of β^*, the solution to limiting score function

$$U(\beta^*) = \int_0^{\infty} \left(\mathscr{E}(\sum_{i=1}^{n} Z_i d\bar{N}_i(t)) - \frac{\mathscr{E}(S^{(1)}(\beta, t))}{\mathscr{E}(S^{(0)}(\beta, t))} \mathscr{E}(\sum_{i=1}^{n} d\bar{N}_i(t)) \right),$$

where the expectation $\mathscr{E}$ is with respect to the true model (e.g. White 1982, Struthers and Kalbfleisch, 1986).

Remark: $\beta^* \neq \beta_0$.

Limiting Treatment Effect: The Independent Case

- Suppose $T_1 \perp T_2|Z$, and $T_k|Z \sim \text{Exp}(\lambda_k e^{\beta_k Z})$.
- If $T = \min(T_1, T_2)$ and $T|Z \sim \text{Exp}(\lambda_1 e^{\beta_1 Z} + \lambda_2 e^{\beta_2 Z})$.
- Random censoring: $C \perp T|Z$, $P(C \geqslant s|Z) = e^{-\rho s}$.
- This independent assumption gives

$$\begin{aligned}\mathscr{E}(Y_i(t)|Z_i) &= \mathscr{E}(Y_i^{\dagger}(t)|Z_i)\mathscr{E}(Y_i^{\ddagger}(t)|Z_i) \\ &= P(T > t|Z_i)P(C > t|Z_i) \\ &= \mathcal{F}(t|Z_i)G(t).\end{aligned}$$

- The expectations in the score equation for β^*, for example $\mathscr{E}(Z_i d\bar{N}_i(t))$, can be obtained as follows

$$\begin{aligned}\mathscr{E}(Z_i d\bar{N}_i(t)) &= \mathscr{E}(Z_i Y_i(t) dN_i(t)) \\ &= \mathscr{E}_{Z_i}\{\mathscr{E}_{Y_i(t)|Z_i}[\mathscr{E}_{dN_i(t)|Y_i(t),Z_i}(Z_i Y_i(t) dN_i(t))]\} \\ &= P(Z_i = 1) f(t|Z_i = 1) G(t) dt.\end{aligned}$$

- $\mathscr{E}(d\bar{N}_i(t))$, $\mathscr{E}(S^{(1)}(\beta, t))$, and $\mathscr{E}(S^{(0)})(\beta, t)$ can be obtained similarly.

- β^* solution is:

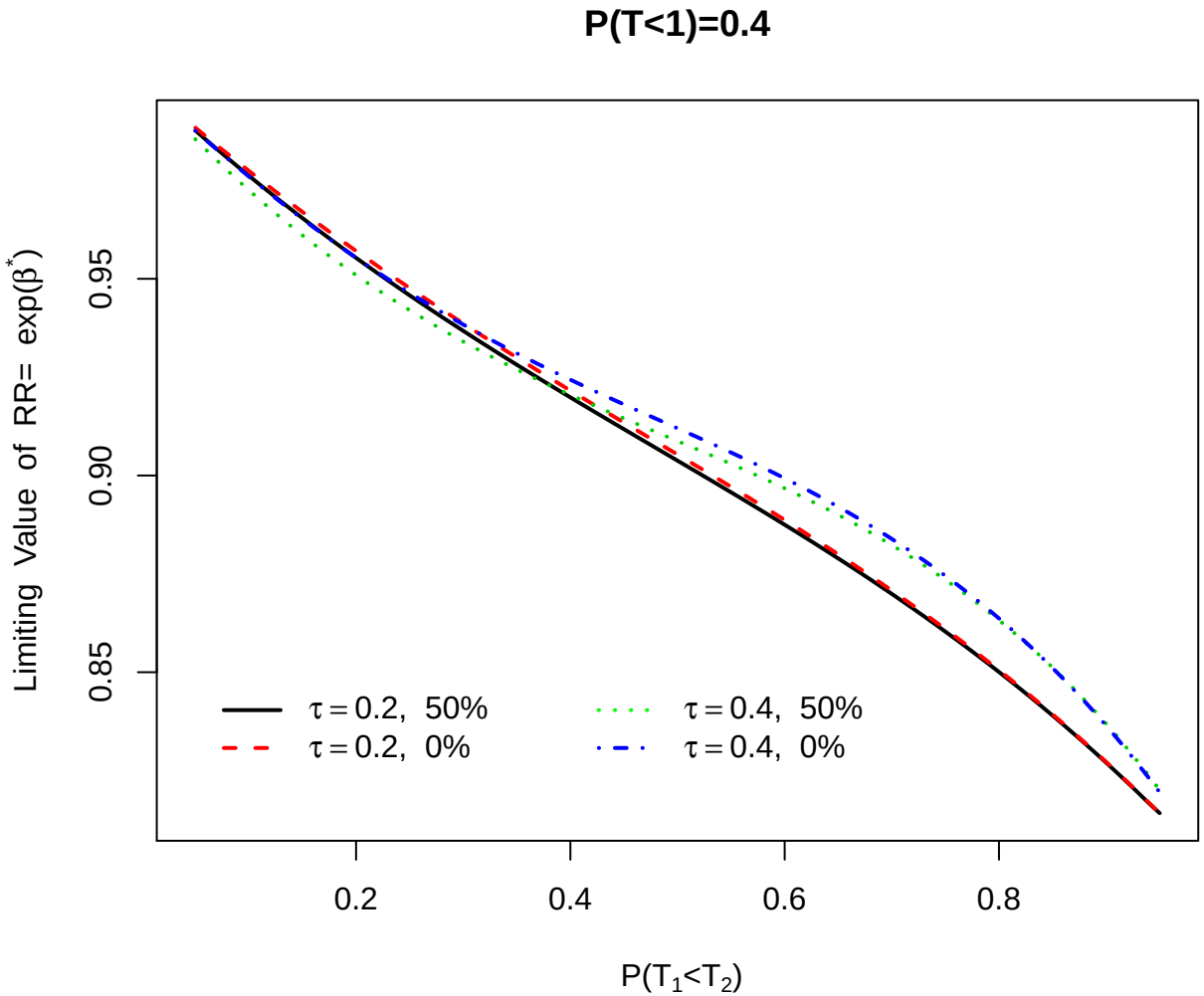
$$e^{\beta^*} = \frac{\lambda_1 e^{\beta_1} + \lambda_2 e^{\beta_2}}{\lambda_1 + \lambda_2}$$
$$= (\frac{\lambda_1}{\lambda_1 + \lambda_2})e^{\beta_1} + (\frac{\lambda_2}{\lambda_1 + \lambda_2})e^{\beta_2}$$

- The solution β^* is "weighted" average of β_1 and β_2.
- The "weights" are determined by λ_1 and λ_2, the frequencies of occurrence of the two type events.

Limiting treatment effect: Associated Components

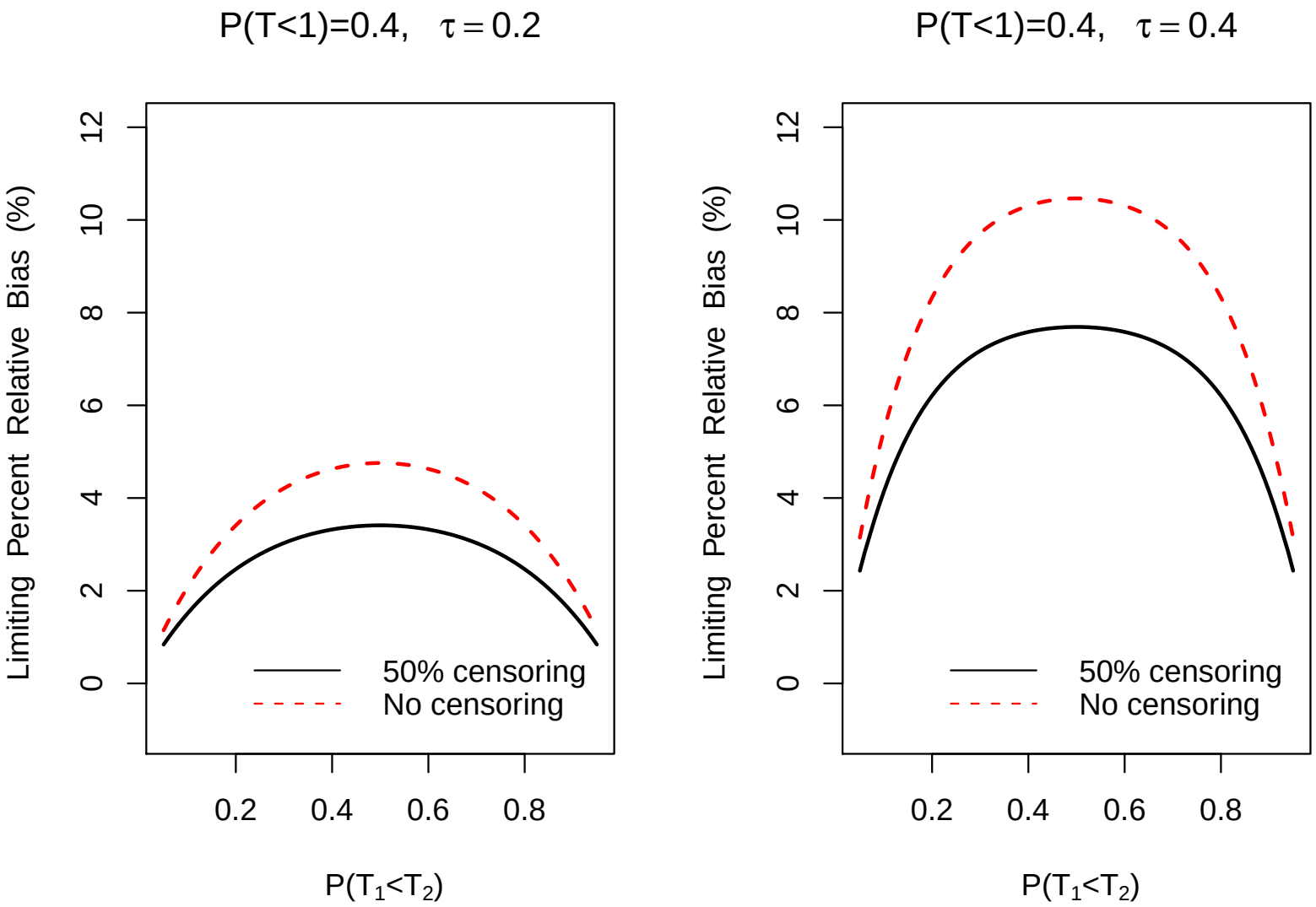
- Use Clayton Copula to model the association.
- The solution of β^* is obtained by numerical integration.
- Some surprising observations on the relation between β^* and β_1, β_2, ϕ.

Limiting Value: Dependent Case with Unequal Treatment Effect



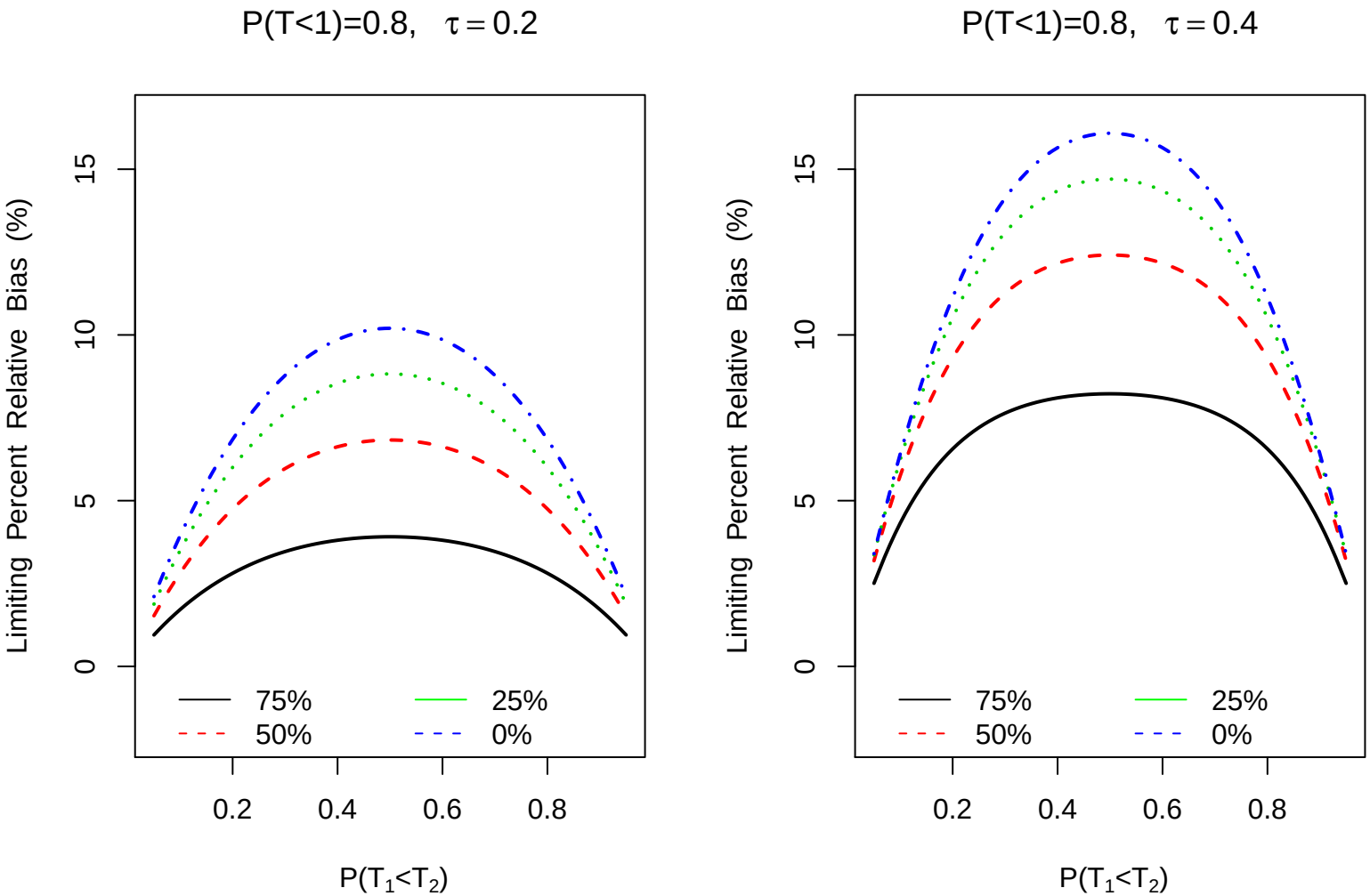
- 60% administrative censoring
- Different random censoring due to withdrawal
- Treatment effects: $\beta_1 = \log(0.8)$ and $\beta_2 = \log(1.0)$

Limiting Value: Dependent Case with Equal Treatment Effect



- 60% administrative censoring
- Different random censoring due to withdrawal
- Treatment effects: $\beta_1 = \beta_2 = \log(0.8)$

Limiting Value: Dependent Case with Equal Treatment Effect



- 20% administrative censoring
- Different random censoring due to withdrawal
- Treatment effects: $\beta_1 = \beta_2 = \log(0.8)$

- Independent components: $T_1 \sim \text{Exp}(\lambda_1)$ and $T_2 \sim \text{Exp}(\lambda_2)$.
- Associated components: $\tau = 0.2, 0.4$.
- Consider both administrative censoring and random censoring.
- D is number of events required for CEP analysis:

$$D = \frac{4(z_{1-\alpha} + z_{1-\gamma})^2}{(\beta^*)^2},$$

  α is type I error for one-sided test;
  $1 - \gamma$ is power;
  β^* is the limiting value of treatment effect estimate;
- The number of subjects required is obtained based on the expected rate of CEP events during the follow-up.

Independent Components With Equal Treatment Effects

$$P(T_1 < T_2|Z = 0) = 0.25, \beta_1 = \log(0.8), \beta_2 = \log(0.8)$$

% admin. cens	% random cens	n	β^*	mean	ESE	AVE SE1	AVE SE2	95%CI	Power
60	60	3,410	-.223	-.223	.090	.090	.090	95—95	80—80
	40	2,265	-.223	-.224	.090	.090	.090	95—95	80—80
	20	1,694	-.223	-.224	.090	.090	.090	95—95	80—80
	0	1,353	-.223	-.224	.090	.090	.090	95—95	80—80
40	60	2,256	-.223	-.225	.090	.090	.090	95—95	81—81
	40	1,494	-.223	-.224	.090	.090	.090	95—95	81—81
	20	1,115	-.223	-.226	.090	.090	.090	95—95	81—81
	0	889	-.223	-.225	.090	.090	.089	95—95	80—80
20	60	1,678	-.223	-.226	.090	.090	.090	95—95	80—80
	40	1,106	-.223	-.227	.089	.090	.090	96—96	81—81
	20	822	-.223	-.224	.088	.090	.090	95—95	81—81
	0	653	-.223	-.223	.090	.090	.090	95—95	80—80

Independent Component With Unequal Treatment Effect

$$P(T_1 < T_2|Z = 0) = 0.25, \beta_1 = \log(0.8), \beta_2 = \log(1.0)$$

% admin cens	% random cens	n	β^*	mean	ESE	AVE SE1	AVE SE2	95%CI	Power
60	60	60,737	-.051	-.052	.020	.021	.021	96—96	82—82
	40	40,450	-.051	-.052	.021	.021	.021	95—95	81—81
	20	30,317	-.051	-.052	.021	.021	.021	95—95	81—81
	0	24,241	-.051	-.052	.021	.021	.021	95—95	82—82
40	60	40,411	-.051	-.052	.021	.021	.021	95—95	80—80
	40	26,892	-.051	-.052	.021	.021	.021	95—95	81—81
	20	20,143	-.051	-.052	.021	.021	.021	95—95	82—82
	0	16,099	-.051	-.052	.021	.021	.021	95—95	80—80
20	60	30,242	-.051	-.052	.021	.021	.021	95—95	80—80
	40	20,101	-.051	-.052	.021	.021	.021	95—95	82—82
	20	15,041	-.051	-.052	.021	.021	.021	94—94	81—81
	0	12,011	-.051	-.052	.021	.021	.021	94—94	81—81

Associated Component With Equal Treatment Effects

$\tau = 0.4$, $P(T_1 < T_2|Z = 0) = 0.25$, $\beta_1 = \log(0.8)$, $\beta_2 = \log(0.8)$

% admin. cens	% random cens	n	β^*	mean	ESE	AVE SE1	AVE SE2	95%CI	Power
60	60	3,825	-.210	-.210	.085	.085	.085	96—96	79—79
	40	2,611	-.207	-.207	.084	.084	.084	95—95	80—80
	20	2,009	-.204	-.204	.086	.082	.082	94—94	79—79
	0	1,619	-.203	-.203	.084	.082	.082	95—95	79—79
40	60	2,653	-.205	-.205	.083	.083	.083	94—94	80—80
	40	1,825	-.201	-.202	.081	.081	.081	95—95	80—80
	20	1,402	-.198	-.199	.079	.080	.080	95—95	81—81
	0	1,140	-.196	-.197	.079	.079	.079	95—95	80—80
20	60	2,033	-.202	-.203	.081	.082	.082	95—95	80—80
	40	1,393	-.198	-.199	.081	.080	.080	95—95	80—80
	20	1,056	-.196	-.197	.079	.079	.079	95—95	81—81
	0	857	-.194	-.194	.077	.078	.078	95—95	80—80

Associated Components With Unequal Treatment Effects

$\tau = 0.4$, $P(T_1 < T_2 | Z = 0) = 0.25$, $\beta_1 = \log(0.8)$, $\beta_2 = \log(1.0)$

% admin. cens	% random cens	n	β^*	mean	ESE	AVE SE1	AVE SE2	95%CI	Power
60	60	42,635	-.061	-.062	.025	.025	.025	96—96	82—82
	40	31,360	-.058	-.059	.024	.023	.023	95—95	81—81
	20	25,186	-.056	-.057	.023	.023	.023	95—95	80—80
	0	20,862	-.055	-.056	.022	.022	.022	95—95	81—81
40	60	32,419	-.057	-.057	.023	.023	.023	95—95	79—79
	40	25,866	-.052	-.052	.021	.021	.021	95—95	81—81
	20	21,787	-.049	-.049	.020	.020	.020	95—95	81—81
	0	18,906	-.047	-.047	.019	.019	.019	95—95	80—80
20	60	29,106	-.052	-.052	.021	.021	.021	95—95	80—80
	40	24,662	-.046	-.046	.018	.019	.019	95—95	79—79
	20	23,192	-.041	-.042	.016	.016	.016	96—96	81—81
	0	21,535	-.038	-.038	.015	.015	.015	96—96	80—80

The Marginal Model of Wei, Lin, and Weissfeld (1989):

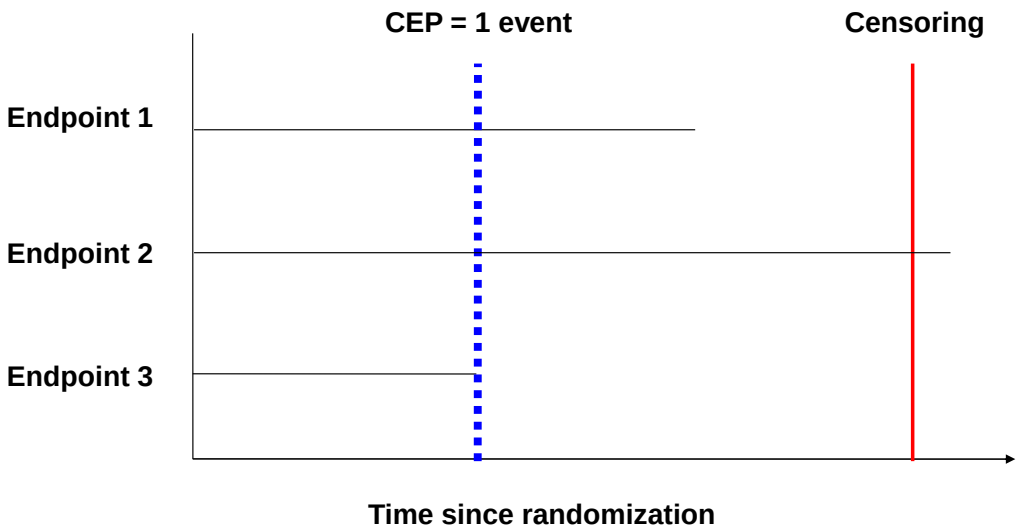
- Model-free to the dependence structure among the multivariate failure times, *i.e.*, components in CEP.
- Fit ordinary Cox model to each component and estimate the regression coefficients.
- Use robust variance estimate in inference to account for possible correlation in the data.

Advantages:

- Can be easily implemented in R or SAS.
- Affords great flexibility in formation of strata and risk sets.
- Well-developed variance estimator—Robust Variance Estimator.

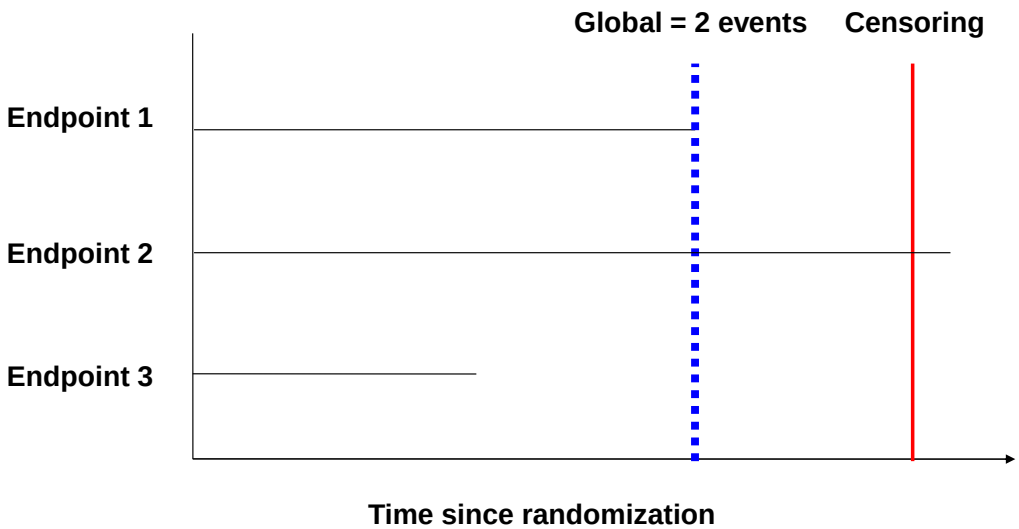
Comparison of Multivariate Analysis and CEP

CEP analysis vs. Global Analysis



Comparison of Multivariate Analysis and CEP

CEP analysis vs. Global Analysis



- Let $\lambda_{ki}(t) = \lambda_{k0}(t)e^{\beta_k Z}$ be the hazard function for kth event.
- The kth event-specific score function for β_k is

$$U_k(\beta_k) = \sum_{i=1}^{n} \int_0^{\infty} \left(Z_i - \frac{S_k^{(1)}(\beta_k, t)}{S_k^{(0)}(\beta_k, t)} \right) d\bar{N}_{ik}(t).$$

  $S_k^{(1)}(\beta, u) = \sum_{i=1}^{n} Y_{ik}(t) Z_i^r \exp\{\beta_k Z_i\}, r = 0, 1.$
- $\bar{N}_i(t) = \{\bar{N}_{ik}(t), k = 1, ..., K\}$ is observed multivariate counting process.
- $Y_{ik}(t)$ is the at risk indicator function for the kth event of subject i.
- Global estimate is weighted average of $\hat{\beta}_k$s.

Empirical Study II: Associated Components With Equal Treatment Effects

$\tau = 0.4$, $P(T_1 < T_2|Z = 0) = 0.25$, $\beta_1 = \log(0.8)$, $\beta_2 = \log(0.8)$

%admin. cens	% random cens	n	β^*	mean	ESE	AVE SE	95%CI	Power
60	60	3,387	-.223	-.224	.091	.09	95	80
	40	2,247	-.223	-.224	.089	.09	95	81
	20	1,679	-.223	-.225	.091	.09	94	81
	0	1,340	-.223	-.224	.09	.089	95	80
40	60	2,239	-.223	-.225	.091	.09	96	80
	40	1,481	-.223	-.224	.088	.089	95	81
	20	1,104	-.223	-.225	.086	.088	96	82
	0	879	-.223	-.225	.085	.088	96	82
20	60	1,666	-.223	-.225	.088	.089	96	82
	40	1,097	-.223	-.225	.088	.088	95	81
	20	815	-.223	-.224	.086	.087	95	83
	0	648	-.223	-.223	.086	.086	95	83

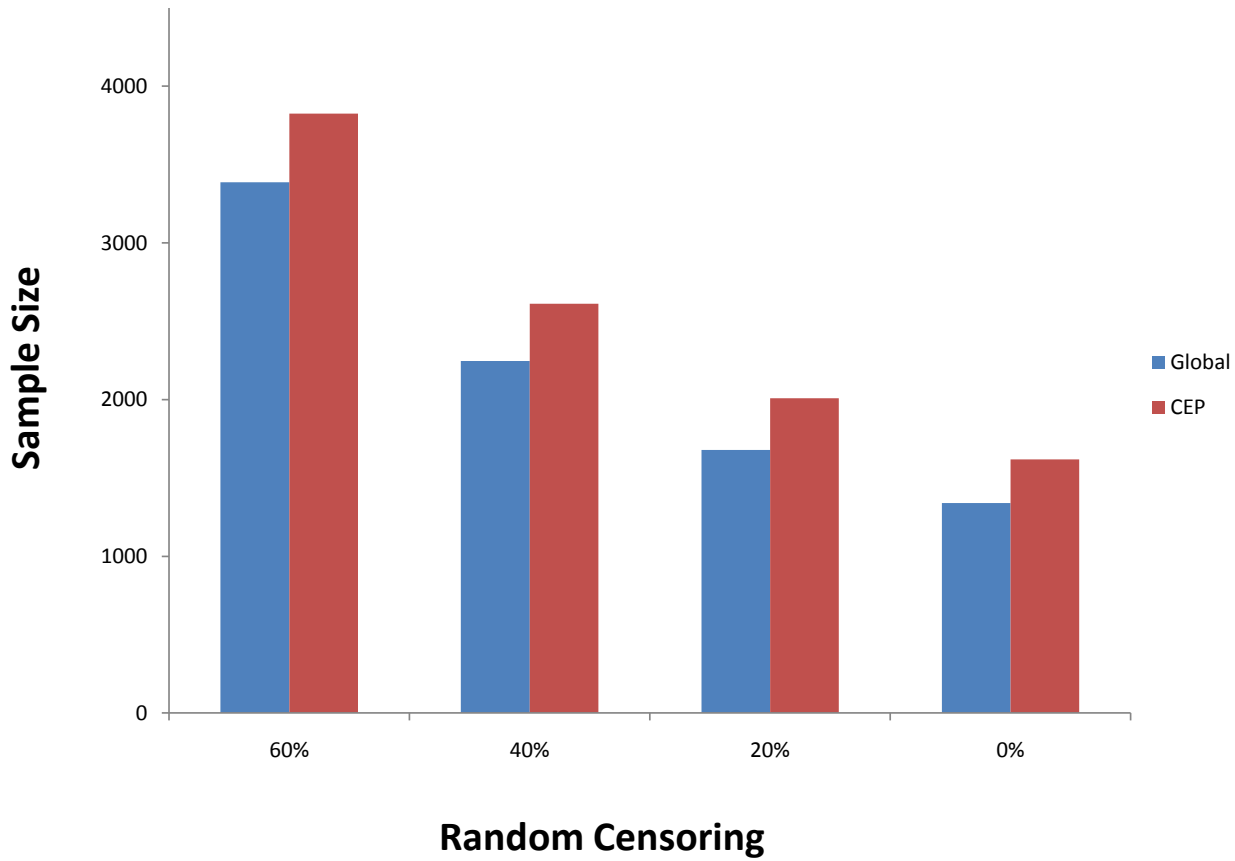
Empirical Study II: Associated Components With Unequal Treatment Effects

$\tau = 0.4$, $P(T_1 < T_2 | Z = 0) = 0.25$, $\beta_1 = \log(0.8)$, $\beta_2 = \log(1.0)$

% admin. cens	% random cens	n	β^*	mean	ESE	AVE SE	95%CI	Power
60	60	35,767	-.067	-.067	.026	.027	95	81
	40	24,218	-.066	-.067	.026	.026	96	81
	20	18,694	-.065	-.067	.026	.026	95	82
	0	14,937	-.065	-.066	.026	.026	94	82
40	60	24,327	-.066	-.066	.026	.026	96	81
	40	16,760	-.065	-.066	.025	.026	95	83
	20	12,771	-.064	-.065	.025	.025	94	83
	0	12,932	-.064	-.065	.025	.025	95	83
20	60	18,514	-.065	-.066	.026	.026	94	82
	40	12,367	-.065	-.066	.026	.025	94	83
	20	9,115	-.065	-.066	.025	.025	95	84
	0	7,182	-.066	-.067	.026	.025	95	84

Sample Size Requirement: Global vs CEP

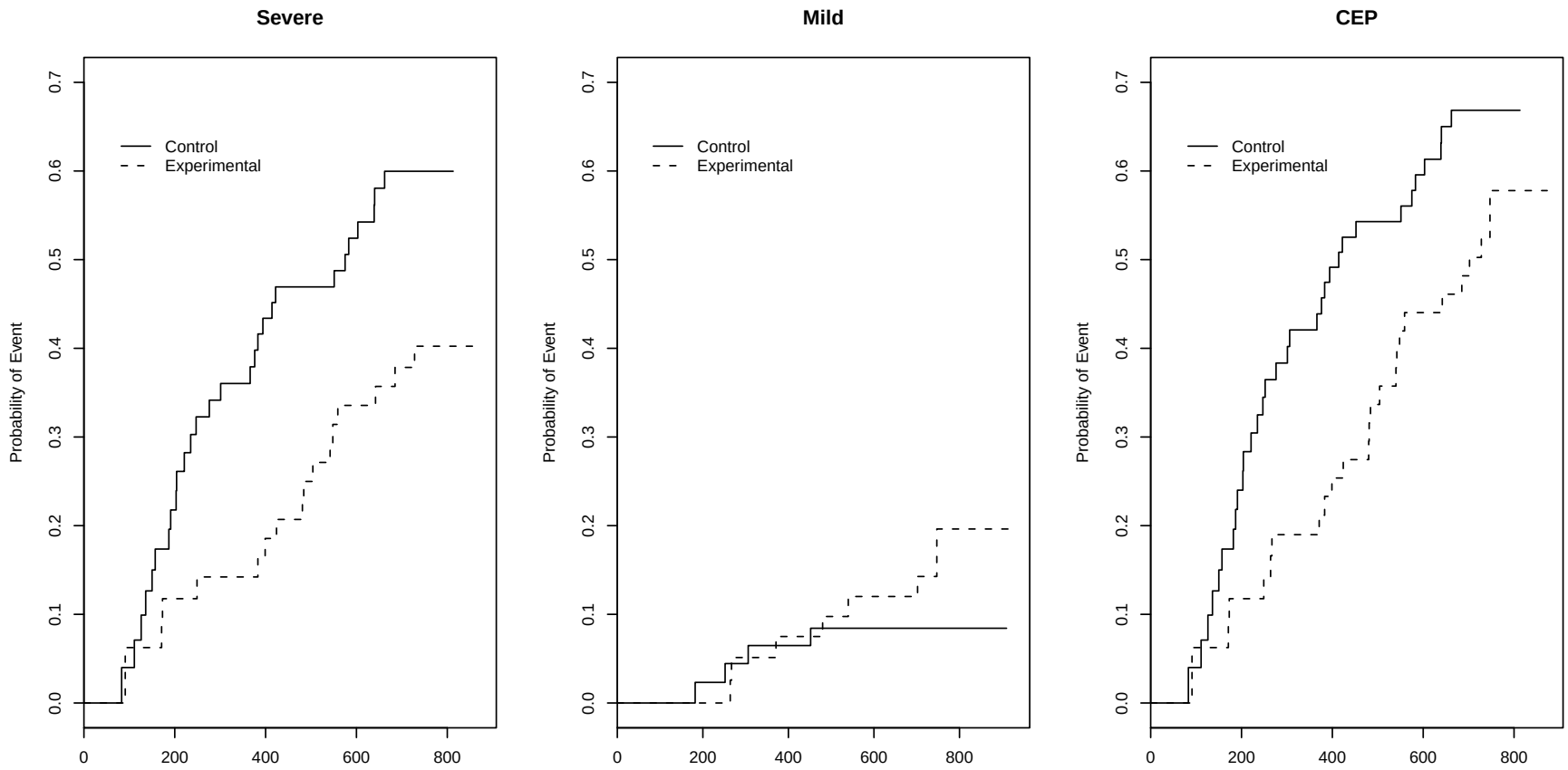
60% admin. cens., $\tau = 0.4$, $P(T_1 < T_2 | Z = 0) = 0.25$,
$\beta_1 = \log(0.8)$, $\beta_2 = \log(0.8)$



A Real Data Example

- An asthma management study—an experimental intervention was tested to delay the time to exacerbation.
- Two endpoints:
  - Endpoint I: severe exacerbation.
  - Endpoint II: mild exacerbation.
- CEP: time to the first event of endpoint I or II.

Probability of events: CEP and its components.



Real Data Analysis: Results

	RR	95%	p-value	p*
Severe	0.53	(0.285, 0.977)	0.042	0.22
Mild	2.14	(0.624, 7.310)	0.227	0.11
CEP	0.665	(0.388, 1.138)	0.137	0.063
Global (WLW)	0.702	(0.405, 1.219)	0.209	

p* is the p-value for testing the proportional hazards assumption.

References to CEP in the Literature

- Search Google Scholar:
  - Search “composite endpoint” and “clinical trial ” —about 2,940 hits since 2000.
  - Search “composite endpoint” and “clinical trial ” and “cardiovascular”— about 2,320 hits since 2000.
- PubMed —about 360 citations in the clinical trial category since 1993.
- Web of Science — about 530 citations:
  - Cardiovascular system and Cardiology (240)
  - Pharmacology and Pharmacy (170)
  - Surgery (81)
  - Hematology (72)
- Search “biostatistics” and “composite endpoint” in “statistics” in Google scholar—about 20 hits.

Some Key References for Composite Endpoints

1. Some uses of Composite Endpoints in clinical trials:
   - 1992. Braunwald *et al.* used a CEP in a cardiovascular trial;
   - 1997. Cannon outlined some rationale for CEP.

2. Critical thinking about CEP in medical literature:
   - 2003. Freemantle et al. raised several cautionary notes;
   - 2005. Montori *et al.*, Neaton *et al.*, and Chi further discussed potential pitfalls and recommendations.

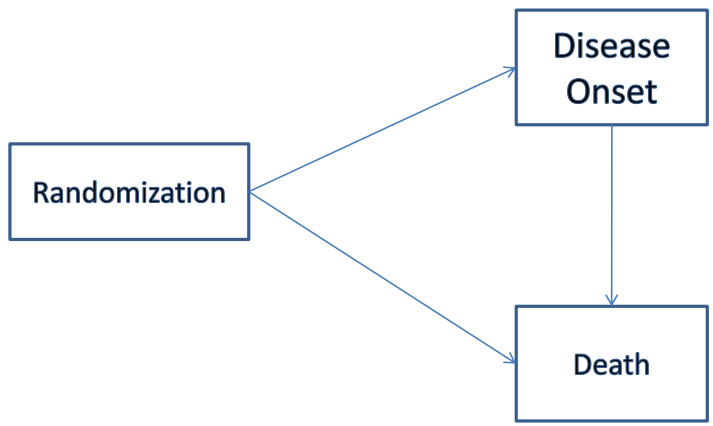
3. Meta-analysis and systematic reviews:
   - 2007-2008. Ignacio Ferreira-Gozalez *et al.*, Bethal *et al.*, Buzney *et al.*, Lim *et al.*, and more.

4. More discussions:
   - 2010. Tomlinson and Detsky; Sheehe.

Rationale Given for Composite Endpoints

- Due to increased event rate, CEP can increase statistical power, reduce sample size and required follow-up time.
- Handles multiplicity problem by using the time-to-first-event.
- Allows the measurement of "overall" benefit of the treatment.
- Useful when a single primary endpoint is hard to choose.
- Can avoid competing risk problem in time-to-event analysis.

Three-state Model



Cannon (1997), Freemantle et al, (2003), Chi (2005), Neaton et al., (2005), Freemantle and Calvert (2007).

Reported Limitations of Composite Endpoints

- Heterogeneity in the treatment effect across components:
  - Poor power for detecting heterogeneity.
  - Interpretation of treatment effect can be difficult.

- Importance of the components may not be equal at the patient level, *e.g.* TIA, stroke, MI, death.

Freemantle et al, (2003), Neaton et al., (2005), Montori et al., (2005), Freemantle and Calvert (2007).

Recommendations Regarding Composite Endpoints

- Individual components should have similar importance to patients;
- Individual components should have similar frequency of occurrence.
- Treatment should have similar effect on all components.

Freemantle et al, (2003), Chi (2005), Neaton et al., (2005), Montori et al., (2005), Freemantle and Calvert (2007), Tomlinson and Detsky (2010).

Recommendations Regarding Composite Endpoints

- Individual components should have similar importance to patients;
- Individual components should have similar frequency of occurrence.
- Treatment should have similar effect on all components.

The last two may not be good recommendations:
Largest bias when components are associated!

Freemantle et al, (2003), Chi (2005), Neaton et al., (2005), Montori et al., (2005), Freemantle and Calvert (2007), Tomlinson and Detsky (2010).

Recommendations Regarding Composite Endpoints

- Individual components should have similar importance to patients;
- Individual components should have similar frequency of occurrence.
- Treatment should have similar effect on all components.

The last two may not be good recommendations:
Largest bias when components are associated!

- Data from all components should be collected until the end of trial.
- Individual components should be analyzed separately as the secondary endpoints.

Freemantle et al, (2003), Chi (2005), Neaton et al., (2005), Montori et al., (2005), Freemantle and Calvert (2007), Tomlinson and Detsky (2010).

Recommendations Regarding Composite Endpoints

- Individual components should have similar importance to patients;
- Individual components should have similar frequency of occurrence.
- Treatment should have similar effect on all components.

The last two may not be good recommendations:
Largest bias when components are associated!

- Data from all components should be collected until the end of trial.
- Individual components should be analyzed separately as the secondary endpoints.

Allow multivariate analysis and facilitate interpretation.

Freemantle et al, (2003), Chi (2005), Neaton et al., (2005), Montori et al., (2005), Freemantle and Calvert (2007), Tomlinson and Detsky (2010).

Discussion

- Cox model for analysis of CEP may not be appropriate: PH does not hold generally.
- Many factors jointly affect the treatment effect estimation:
  - the dependence structure in the individual components;
  - stochastic ordering and occurrence frequencies of individual components;
  - the amount of random censoring;
  - heterogeneity of the treatment effect across the individual components.
- "Equal treatment effect and equal frequency" of individual component may not be valid recommendation.
- The multivariate approach generally outperforms the CEP:
  - provides the average effect of treatment—facilitates the interpretation;
  - can achieve higher power and accuracy;
  - can claim treatment effect on individual components—permit intent-to-treat analysis.

Acknowledgment

- CANNeCTIN Graduate Student Award;
- NSERC Postgraduate Scholarship (PGS);
- Ontario Graduate Scholarship (OGS).

Reference

Bethel MA *et al.* (2008) Determining the most appropriate components for a composite clinical trial outcome. *American Heart Journal* doi:10.1016/j.ahj.2008.05.018.

Buzney EA, Kimball AB. (2008) A critical assessment of composite and coprimary endpoints: a complex problem. *J Am Acad Dermatol* 59:890-6.

Braunwald E, Cannon CP, McCabe CH, (1992) An approach to evaluating thrombolytic therapy in acute myocardial infarction. The 'unstatisfactory outcome' end point. *Circulation* 86:683-687.

Braunwald E, Cannon CP, McCabe CH. (1993) Use of composite endpoints in thrombolysis trials of acute myocardial infarction *The American Journal of Cardiology* 72 (19), pg. G3-G12

Cannon CP (1997) Clinical perspectives on the use of composite endpoints. *Controlled clinical trials* 18:517-529.

Chi GYH (2005) Some issues with composite endpoints in clinical trials. *Fundamental & Clinical Pharmacology* 19:609-619.

Cook, R.J. and Lawless, J.F. (2007). The Statistical Analysis of Recurrent Events. Springer, New York.

Cox DR (1972) Regression models with lifetables (with discussion). *J Roy Stat Soc Ser B* **34,** 187-220

Ferreira-González *et al.* (2007) Problems with use of composite end points in cardiovascular trials:systematic review of randomised controlled trials. *BMJ* doi:10.1136

Ferreira-Gonzalez I, Permanyer-Miralda G, Busse JW, Bryant DM, Montori VM, Alonso-Coello P, et al. (2007) Methodologic discussions for using and interpreting composite endpoints are limited, but still identify major concerns.*J Clin Epidemiol* 60:651e7.

Ferreira-González *et al.* (2007) Composite endpoints in clinical trials: the trees and the forest. *J Clin Epidemiol* 60:660-661.

Ferreira-González *et al.* (2008) Composite enpoints in clinical trials. *Rev Esp Cardiol* 61(3):283-90.

Freemantle N *et al.* (2003) Composite outcomes in randomized trials: greater precision but with greater uncertainty? *JAMA* Vol 289. No.19 2545-2575.t

Freemantle N, Calvert M. (2007) Weighing the pros and cons for composite outcomes in clinical trials. *Journal of Clinical Epidemiology* 60:658-659

Reference

Lim E*et al.* (2008) Composite outcomes in cardiovascular research: a survey of randomized trials. *Ann of Intern Med* 149:612-617.

Montori VM, *et al.*, (2005) Validity of composite end points in clinical trials. *BMJ* Vol 330 2005.

Neaton JD, Gray G, Zuckerman BD, Konstam M, (2005) Key issues in end point selection for heart failure trials: composite end points. *Journal of Cardiac Failure* Vol.11 No. 8 2005.RTG

Prentice, RL. and Cai J. (1992) Covariance and survivor function estimation using censored multivariate failure time data. *Biometrika* 79, 3:495-512.

Struthers CA and Kalbfleish JD (1986) Misspecified Proportional Hazard Models. *Biometrika*, 73, 363-369.

Wei, LJ, Lin,DY, and Weissfeld L. (1989) Regression Analysis of Multivariate Incomplete Failure Time Data by Modeling Marginal Distributions. *Journal of American Statistical Association* Vol. 84 1065-1073.

White H. (1982) Maximum Likelihood Estimation of Misspecified Models. *Econometrica* Vol 30:1-25.

Yusuf, *et al.* (2000) Effects of An Angiotensin-Converting-Enzyme Inhibitor, Ramipril, on Cardiovascular Events in High-Risk Patients. *The New England Journal of Medicine* Vol 342 No. 3 145-153.

Thank you!