

**Statistical Methods for the Evaluation of
Treatment Effectiveness in the Presence of
Competing Risks**

Ravi Varadhan

Assistant Professor

(On Behalf of the Johns Hopkins DEcIDE Center)

Center on Aging and Health

Johns Hopkins University

Presented at the AHRQ Symposium

June 02, 2009

Competing Risks

The task of estimating the likelihood that a patient will benefit from an intervention for a specific condition, *in the presence of other conditions*, is a **competing risks** problem.

Treatment Efficacy in RCTs: An Example

- Green and Byar (1980): Efficacy of Diethylstilbestrol (DES) for treating prostate cancer
- Death from prostate cancer was primary outcome
- Death from cardiovascular and other causes were competing events
- $N = 483$; 344 deaths over a 6-year followup
- $PC = 149$; $CVD = 139$; $Other = 56$
- Prognostic variables: treatment indicator, age (AG), weight index (WT), performance rating (PF), history of heart disease (HX), serum hemoglobin (HG), size of primary lesion (SZ), and Gleason stage (SG)

Treatment Efficacy in RCTs: An Example

- Three key questions:
 1. Does the treatment reduce the rate of prostate cancer death?
 2. Does the treatment reduce absolute risk of prostate cancer death?
 3. Which groups of men are likely to benefit from (or be harmed by) DES?
- Q1 deals with the average Tx effect
- Q2 deals with the overall absolute risk reduction due to Tx
- Q3 deal with real-world benefit for an individual X

Some Basic Notation

- Data: $\{Y_i, \Delta_i, \Delta_i * \epsilon_i, X_i\}$
 - Y_i = follow-up time
 - $\Delta_i = \begin{cases} 1 & ; \text{ if failed} \\ 0 & ; \text{ if right-censored} \end{cases}$
 - $\epsilon_i \in \{1, 2, \dots, K\}$ = cause of failure
 - X_i = vector of covariates, including Tx indicator
- Problem formulation in terms of ‘observables’

Statistical Methods for *Univariate* Survival Analysis

- Triumvirate of survival analysis: Kaplan-Meier plot, log-rank test, and Cox PH regression
- **One-sample:** Kaplan-Meier (product-limit estimator) survival plots
- **K-sample:** Peto's log-rank test for comparing hazards or survival between K groups
- **Regression:** Cox's proportional hazards model
- These methods require special attention when applied to competing risk problems

Different Estimands in Competing Risks

- An estimand is a theoretical quantity that is of inferential interest (e.g. mean, odds ratio, rates)
- Three important estimands in CR problems:
 - Cause-specific hazard (CSH)
 - Cumulative incidence function (CIF)
 - Event-free survival distribution (EFS)
- We will define and discuss this for two types of events: primary (type1) and competing (type2)

Basic Definitions

- Cause-specific hazard: (CSH):

$$\lambda_1(t|X) dt := \Pr(t < Y \leq t + dt, K = 1 | Y \geq t)$$

- Instantaneous **rate** of event at time t (Speedometer reading)

- Cumulative incidence function (CIF):

$$F_1(t|X) := \Pr(Y \leq t, K = 1)$$

- **Probability** of event upto time t (Odometer reading!)

- Event-free survival distribution (EFS):

$$S(t|X) := \Pr(Y \geq t)$$

$$\lambda_{12}(t|X) dt := \Pr(t < Y \leq t + dt | Y \geq t) = \lambda_1(t|X) + \lambda_2(t|X)$$

- Probability of surviving without any events until time t

Cause-Specific Hazard

- Cause-specific hazard: (CSH):

$$\lambda_1(t|X) dt := \Pr(t < Y \leq t + dt, K = 1 | Y \geq t)$$

- Easiest estimand to model
- Competing event is treated as right-censoring
- Cox PH or RR model will work for regression
- Peto's log-rank test will work for K-sample testing
- However, $1 - \exp\{-\Gamma_1(t|X)\}$ does not have valid interpretation as failure probability
- It over-estimates cumulative failure probability of Type1 events

Cumulative Incidence Function

- Cumulative incidence function (CIF):

$$F_1(t|X) := \Pr(Y \leq t, K = 1)$$

$$F_1(t|X) = \int_0^t \lambda_1(u|X)S(u|X)du$$

Where: $S(u|X) = \exp[-\int_0^u \lambda_{12}(v|X)dv]$ is the overall survival function

- Extension of Peto's log-rank test (Gray 1988) for testing equality of CIF curves
- Gail's absolute cause-specific risk (e.g. 5-year absolute risk of breast cancer; Benichou and Gail, 1995)
- Direct regression modeling: Fine & Gray (1999)

Event-Free Survival

- Event-free survival distribution (EFS):

$$S(t|X) := \Pr(Y \geq t)$$

$$\lambda_{12}(t|X) dt := \Pr(t < Y \leq t+dt | Y \geq t) = \lambda_1(t|X) dt + \lambda_2(t|X) dt$$

- Obviously depends on both primary and competing event rates
- Analysis of time to first event of any type
- A main advantage: it reduces CR problem to univariate survival analysis
- Easily modelled using standard Cox models, log-rank tests, and 1 - Kaplan-Meier estimator is valid

DES for Prostate-Cancer - Example

- Green and Byar (1980): Efficacy of Diethylstilbestrol (DES) for treating prostate cancer
- Death from prostate cancer was primary outcome
- Death from cardiovascular and other causes were competing events

Treatment Efficacy in RCTs: An Example

- Three key questions:
 1. Does the treatment reduce the rate of prostate cancer death?
 2. Does the treatment reduce absolute risk of prostate cancer death?
 3. Which groups of men are likely to benefit from (or be harmed by) DES?

Table 1
 Parameter estimates for event-free survival and cause-specific hazards.
 Standard errors are given in parenthesis

Covariate	Event-free Survival	Cause-specific hazard			Cumulative Incidence		
		Cancer	CVD	Other	Cancer	CVD	Other
Treatment	-0.19 (0.11)	-0.55 (0.17)	0.36 (0.17)	-0.58 (0.28)	-0.43 (0.17)	0.38 (0.17)	-0.65 (0.29)
AG	0.29 (0.09)	0.00 (0.14)	0.34 (0.13)	0.77 (0.20)	-0.11 (0.15)	0.17 (0.13)	0.62 (0.24)
WT	0.20 (0.09)	0.19 (0.14)	0.04 (0.15)	0.53 (0.22)	0.09 (0.15)	-0.03 (0.15)	0.44 (0.21)
PF	0.41 (0.17)	0.26 (0.26)	0.48 (0.27)	0.54 (0.42)	0.14 (0.26)	0.32 (0.28)	0.14 (0.44)
HX	0.44 (0.11)	-0.10 (0.18)	1.15 (0.19)	0.02 (0.29)	-0.25 (0.18)	1.14 (0.19)	-0.27 (0.29)
HG	0.30 (0.12)	0.47 (0.18)	0.02 (0.20)	0.36 (0.30)	0.32 (0.19)	-0.21 (0.21)	0.16 (0.29)
SZ	0.68 (0.16)	1.16 (0.20)	-0.22 (0.36)	0.72 (0.42)	0.84 (0.21)	-0.53 (0.37)	0.27 (0.44)
SG	0.40 (0.12)	1.35 (0.20)	-0.02 (0.19)	-0.45 (0.30)	1.30 (0.20)	-0.22 (0.19)	-0.68 (0.30)

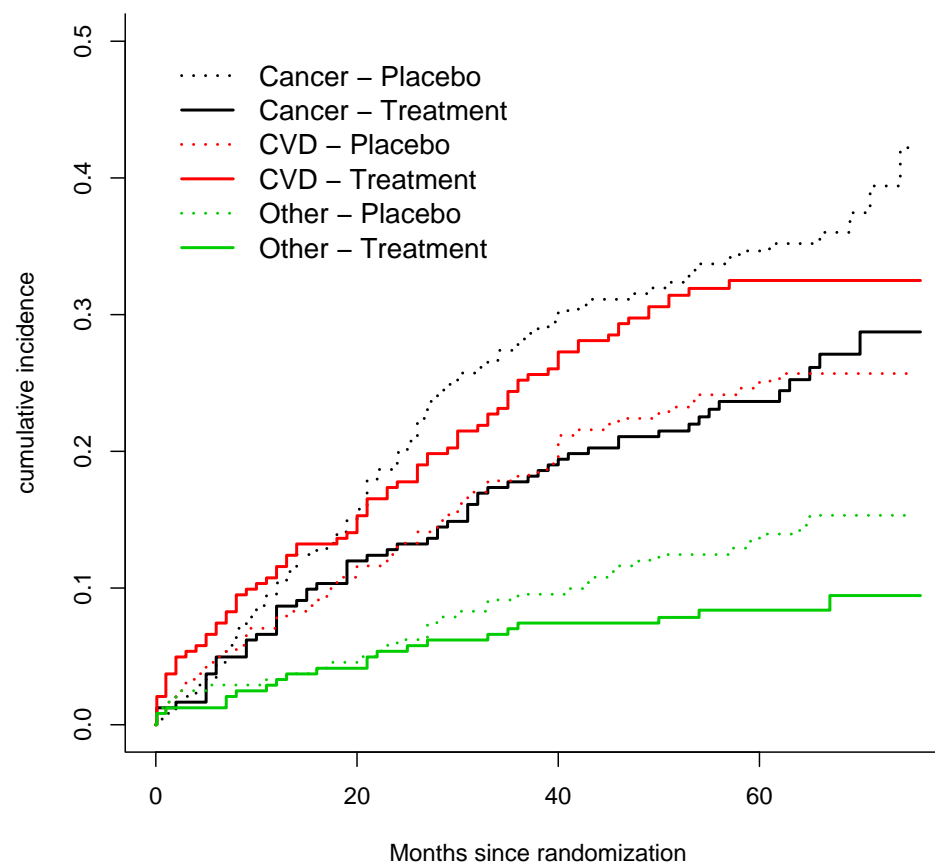


Figure 1:

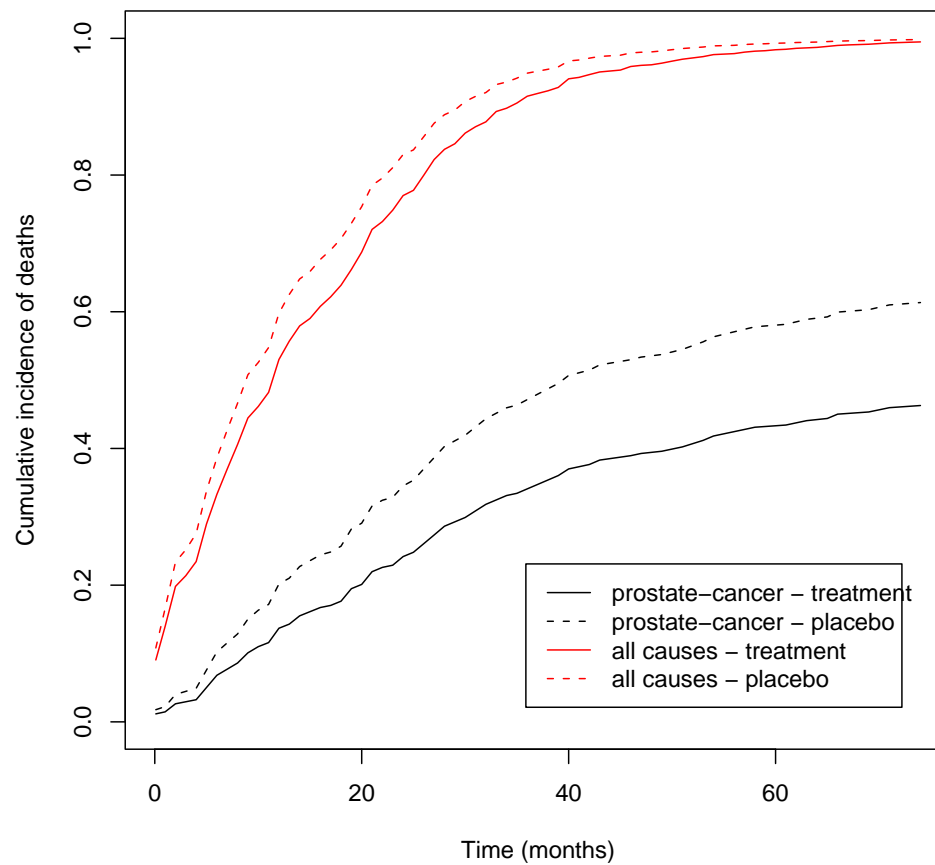


Figure 2: Evaluation of overall treatment benefit for an individual with history of heart disease (HX=1), large size primary lesions ((SZ=1), and advance Gleason stage (SG=1)

Which Estimand to Use?

- Depends on the objective
- Etiology - identification of risk factors
- 5-year risk prediction (e.g. Framingham risk score)
- Average Tx effect estimation
- Tx risk-benefit assessment for the individual

When is CSH appropriate?

- Models the intensity of the underlying event-generating process
- CSH can capture biological effects (“efficacy”)
- CSH is useful for average Tx effect estimation
- Useful for risk factor etiology (Cox models in Epi studies)
- When assessing Tx effects: all CSHs should be reported

When is EFS appropriate?

- Frequently used in RCTs to estimate Tx effect (composite outcome or disease-free survival)
- Increased power if Tx effect on primary and competing events has the same sign and similar magnitude
- Represents overall beneficial effect of Tx
- Decreased power if Tx effects are widely different or have opposite signs
- Validity and interpretation is problematic
- When assessing Tx effects: 1 - EFS should be reported

When is CIF appropriate?

- CSH cannot be used to estimate absolute risk
- $1 - \exp\{-\Gamma_1(t|X)\}$ overestimates absolute risk
- Framingham risk score and many other risk prediction models use this incorrect approach
- CIF is the right approach for risk prediction
- Risk-benefit assessment of Tx (comparative evaluation of CIFs for good and adverse outcomes)
- When assessing Tx effects: CIF for primary cause should be reported

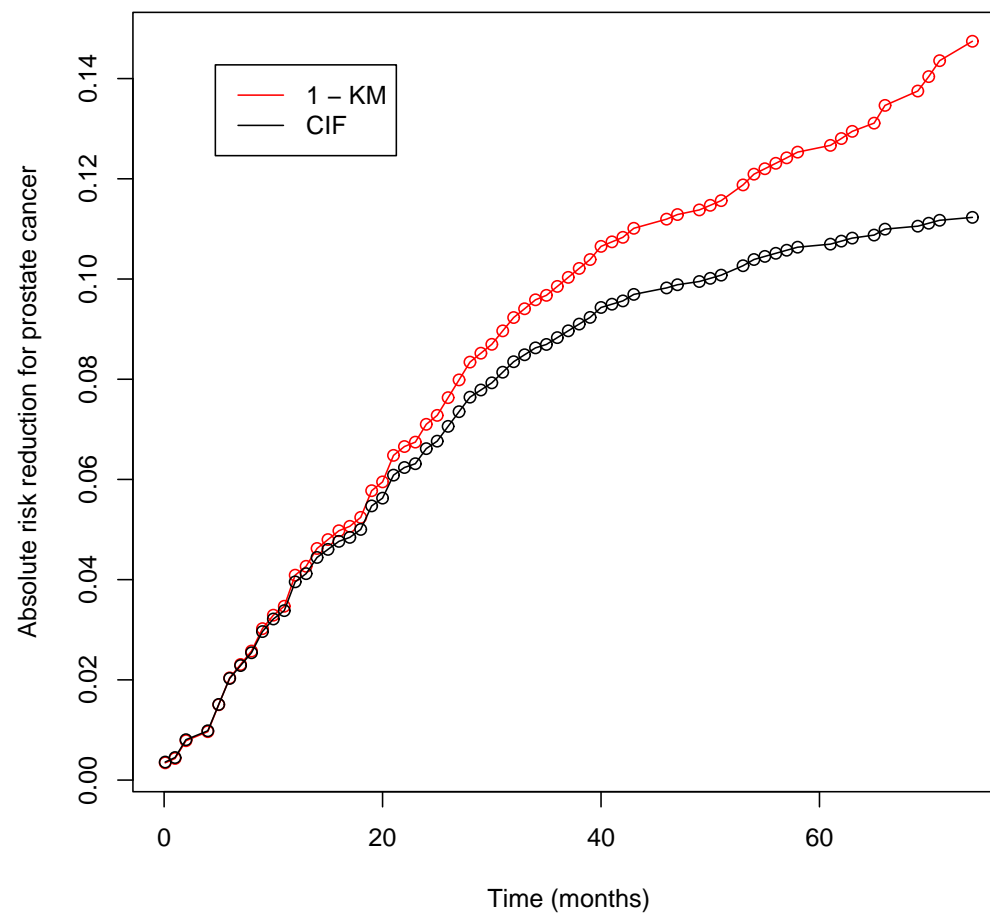


Figure 3:

Conclusions

- Competing risks issues important in evaluation of Tx effectiveness
- Three major estimands - estimable using observed data
- Which one to use depends on the objective
- Recommend a comprehensive approach in reporting RCT results
- 1 - KM should not be used for absolute risk and absolute risk reduction
- Software implementing competing risks methodology is becoming available

Acknowledgements

Jodi Segal, M.D., M.P.H.

Carlos Weiss, M.D., M.H.S.

Cynthia Boyd, M.D., M.P.H.

Christine Weston, Ph.D.

Dan Scharfstein, , Sc.D.

Properties of Estimands: A Simulation Example

- Two event types 1 and 2; two treatment groups A and B
- Generated data from four (uncorrelated) exponential distributions
- Testing equality of CSHs:

$$H_1 : \lambda_1^A(t) = \lambda_1^B(t) \quad (\text{log-rank test})$$

- Testing equality of EFS:

$$S_0 : \lambda_{12}^A(t) = \lambda_{12}^B(t) \quad \text{or} \quad S^A(t) = S^B(t) \quad (\text{log-rank test})$$

- Testing equality of CIFs:

$$I_1 : F_1^A(t) = F_1^B(t) \quad (\text{Gray's K-sample test})$$

Table 2: Empirical rejection probabilities at two-sided 0.05 level (no censoring; sample size is 100 in each group; 10 000 replications). Type 1 is the primary outcome and Type 2 is the competing outcome. Treatment A is a new treatment, and B is either placebo or a standard treatment.

Scenarios	Hypothesis Tests		
	CSH (T^{H_1})	CIF (T^{I_1})	EFS (T^{S_0})
(A) Treatment has no effect on primary event and			
1. No effect on competing event $\lambda_1^A = 1, \lambda_2^A = 1$	0.055	0.051	0.052
2. Decreases competing event rate $\lambda_1^A = 1, \lambda_2^A = 0.6$	0.053	0.25	0.34
3. Increases competing event rate $\lambda_1^A = 1, \lambda_2^A = 1.6$	0.050	0.24	0.45
	CSH (T^{H_1})	CIF (T^{I_1})	EFS (T^{S_0})
(B) Treatment has no effect on competing event and			
1. Decreases primary event rate $\lambda_1^A = 0.6, \lambda_2^A = 1$	0.65	0.50	0.34
2. Increases primary event rate $\lambda_1^A = 1.6, \lambda_2^A = 1$	0.69	0.50	0.45
	CSH (T^{H_1})	CIF (T^{I_1})	EFS (T^{S_0})
(C) Treatment affects both events and			
1. Decreases primary and competing event rates $\lambda_1^A = 0.6, \lambda_2^A = 0.6$	0.70	0.11	0.95
2. Increases primary and competing event rates $\lambda_1^A = 1.6, \lambda_2^A = 1.6$	0.63	0.10	0.91
3. Decreases primary and increases competing event rates $\lambda_1^A = 0.6, \lambda_2^A = 1.6$	0.58	0.89	0.10
4. Increases primary and decreases competing event rates $\lambda_1^A = 1.6, \lambda_2^A = 0.6$	0.73	0.91	0.10

NOTE: $\lambda_1^B = \lambda_2^B = 1$

