Cost Analysis with Censored Data

Yijian (Eugene) Huang

Department of Biostatistics, Emory University

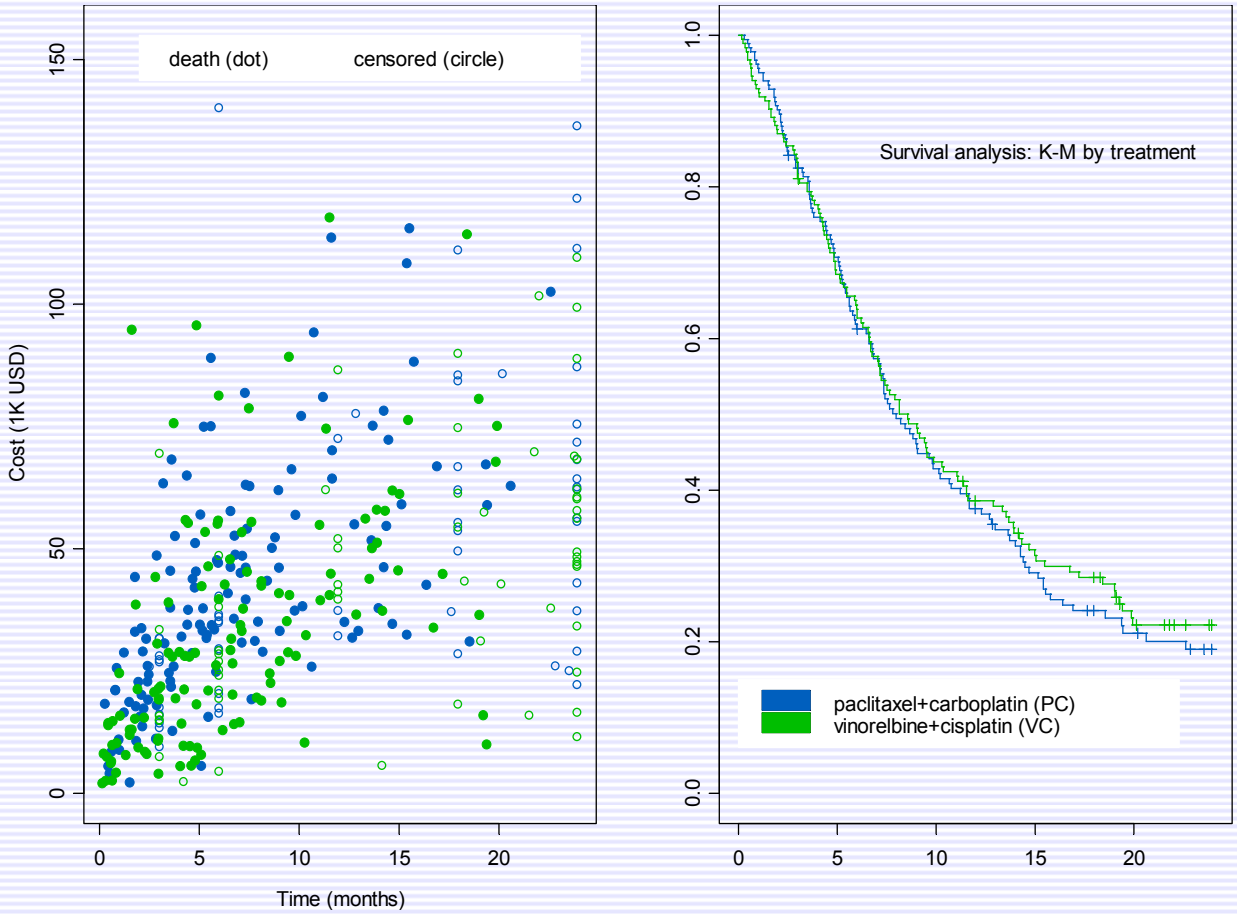
December, 2007

Outline

- Right censoring is common in data from clinical trials and observational studies
- Statistical issues
  1. Induced dependent censoring
  2. Marginal identifiability
- Analysis strategies
  1. Imposing time limit
  2. Joint distribution / modeling with survival time
- Summary and discussion

Example: SWOG lung cancer trial

phase III on advance non-small cell lung cancer (Kelly et al. 2001): size 408



secondary endpoint: cost comparison PC vs. VC?

Issue #1: Induced dependent censoring

T: survival time; U: lifetime cost; C: censoring time

Time scale:

$$X = \min(T, C) \qquad \Delta = I(T \leq C)$$

Cost scale (assuming time-constant cost accumulation rate r):

$$Y = rX = \min(U=rT, rC)$$

$$r \uparrow \Rightarrow \text{cost accumulated at death} \uparrow, \text{ at censoring time} \uparrow$$

Implication: standard survival analysis techniques not applicable to cost-to-event

Issue #2: Marginal identifiability of cost distribution

Of interest: cost-to-event, or lifetime cost, U

Q: Possible to estimate (marginal) distribution of U?

Participants who survive beyond the study duration:

some accumulating little cost during the study

$\Rightarrow$ little info on their cost distribution

$\Rightarrow$ $\Pr(U \leq u)$ not identifiable for any $u \in (0, \infty)$

What can one say, if any, about cost then?

Strategy #1: Imposing time limit

Time-restricted cost:

2-yr-restricted cost = cost accumulated up to min(2 yr, lifetime)

time limit $\leq$ study duration $\Rightarrow$ identifiability

compromise between identifiability and cost of interest

One-sample problem

- Lin et al (1997)
  - partition time span to small intervals
  - mean cost in a small interval = survival rate $\times$ mean cost of alive
  - sum over all intervals
- Bang & Tsiatis (2000) suggested similar and improved estimators

Strategy #1: Imposing time limit - cont'd

Two-sample problem

- one-sample estimation procedure may be used to construct two-sample test, e.g. Ramsey et al (2002)

Regression problem

- Lin (2000): Using inverse probability weighted (IPW) estimation in linear regression to account for censoring

Strategy #1: Imposing time limit - cont'd

Analysis results for the lung cancer trial (Ramsey et al 2002)

	VC	PC	*P-val*
Mean 2 yr-restricted cost	$40,292	$48,940	.004
95% CI	36,226 – 44,359	44,674 – 53,208	

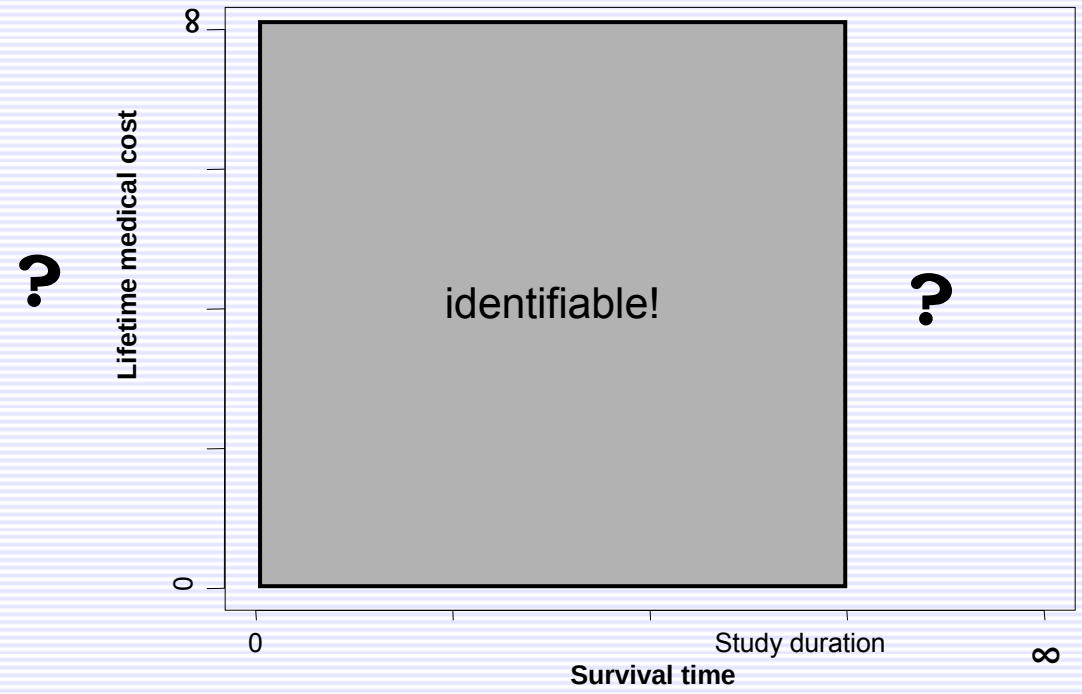
Comments:

- Widely used in practice - censoring can be taken into account by IPW, and so standard software might be used
- Time limit is artificial. A tx favored in time-restricted cost ≠ favored in lifetime cost
- Can be misleading, particularly when tx has an effect on survival time

Strategy #2: Joint distribution with survival time

Joint distribution of (U,T) is largely identifiable with a general data structure (marked point process):

$$X = \min(T, C) \qquad \Delta = I(T \leq C) \qquad Y = U\, I(T \leq C)$$



Strategy #2: Joint distribution with survival time - cont'd

One-sample problem

nonparametric approach (Huang & Louis 1998)

- NPMLE for $\Pr(T \leq t, U \leq u)$
- generalization of K-M estimator

semiparametric approach (Huang & Berry 2006)

- postulate the association structure of (U,T), but leave the marginal distributions of U and T unspecified
- consistently estimate the marginal distribution of U

Strategy #2: Joint distribution with survival time - cont'd

Two-sample problem (Huang & Lovato 2002)

- motivation: no tx effect on T $\Rightarrow$ compare $(U^{(1)}, T^{(1)})$ and $(U^{(2)}, T^{(2)})$
- calibrating tx effect on survival time with, say, $T^{(1)} = \beta\, T^{(2)}$

  compare $(U^{(1)}, T^{(1)})$ and $(U^{(2)}, \beta\, T^{(2)})$

Regression (Huang 2002)

calibration regression: $$\log\begin{pmatrix} T \\ U \end{pmatrix} = \begin{pmatrix} \beta'_0 \\ \beta'_1 \end{pmatrix} Z + \varepsilon$$

- a generalization of accelerated failure time (AFT) model
- simultaneous inference of covariate effects on U and T

Strategy #2: Joint distribution with survival time - cont'd

Analysis results for the lung cancer trial (Huang 2002)

	Survival time			Lifetime cost		
	PC (vs. VC)	LDH	Age	PC (vs. VC)	LDH	Age
Reg coef	.0221	.6335	-.0058	.3400	.1418	-.0050
SE	.1364	.1507	.0065	.1094	.1154	.0059

Comments:

- target lifetime cost U
- difficult to take advantage of cost accumulation data if available
- consider U where end-study survival rate is ~50%?

  Desirable? Yes! Realistic? Maybe not.

Summary and discussion

- For cost analysis in a clinical trial, targeting time-restricted cost would be reasonable if the treatment has little effect on survival time.
- If survival rate is fairly small, say $\leq$ 20%, one might consider joint modeling of lifetime medical cost and survival time. In this case, model assumption can be reasonably checked.
- However, for many trials, treatment has an effect on survival time and survival rate is ~50% or higher at the end of study. What is a sensible yet estimable cost outcome to look at?
- Cost-effectiveness analysis?
- Some of the issues may be more economic than statistical.

References

Bang H & Tsiatis AA (2000) Estimating medical costs with censored data. *Biometrika* 87, 329-43.

Huang Y (2002) Calibration regression of censored lifetime medical cost. *JASA* 97, 318-27. correction, 97, 661.

Huang Y & Berry K (2006) Semiparametric estimation of marginal mark distribution. *Biometrika* 93, 895-910.

Huang Y & Louis TA (1998) Nonparametric estimation of the joint distribution of survival time and mark variables. *Biometrika* 85, 785-98.

Huang Y & Lovato L (2002) Tests for lifetime utility or cost via calibrating survival time. *Statistica Sinica* 12, 707-23.

Kelly K et al (2001) Randomized phase III trial of Paclitaxel plus Carboplatin versus Vinorelbine plus Cisplatin in the treatment of patients with advanced non-small-cell lung cancer: A Southwest Oncology Group trial. *J Clin Oncol* 19, 3210-8.

References – cont'd

Lin DY (2000) Linear regression analysis of censored medical costs. *Biostatistics* 1, 35-47.

Lin DY, Feuer EJ, Etzioni R & Wax Y (1997) Estimating medical costs from incomplete follow-up data. *Biometrics* 53, 419-34.

Ramsey SD et al (2002) Economic analysis of Vinorelbine plus Cisplatin versus Paclitaxel plus Carboplatin for advanced non-small-cell lung cancer. *JNCI* 94, 291-7.