

# Outline

- Issues with truncation by death in clinical trials
- A real-world example
- Definition of causal parameters of interest
- Model identifiability
- Estimation
- Application to the real-world example

# Truncation by death

- Longitudinal studies on outcomes, such as morbidity, health status, and health-related quality of life, allow the assessment of health trajectories over time.
- During follow-ups, however, subjects may be absent for certain scheduled visits or even die before the end of the study.
- Outcomes truncated by death are particularly common in the study population of elderly adults or advanced-stage disease patients. In medical studies, there are many situations when the final outcomes are truncated by death.

## A motivating example

- From Oct 1999 to Jan 2003, Southwest Oncology Group (SWOG) conducted a randomized phase III trial to prospectively compare the treatment of docetaxel and estramustine (DE) with mitoxantrone and prednisone (MP) in patients with metastatic, androgen-independent prostate cancer.
- A previously published result, by Petrylak et al. (2004), has demonstrated an improvement in median survival of nearly two months with DE, as compared to MP treatment.

## The example, continued

- In the trial, enrolled patients were randomly assigned to arm 1 (DE) or arm 2 (MP) for cycles of 21 days. The assigned treatment continued either until disease progression or toxicity occurred or until the maximum of 12 cycles were administered.
- Each patient's quality of life (QOL) was assessed by the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30 (QLQ-C30), administered only at baseline, cycle 4 (the third month), cycle 8 (the sixth month) and the first year.
- QOL-C30: five functional scales, three symptom scales, and a global health and quality of life scale (GQOL).
- We are interested in assessing the effect of DE versus MP on QGOL at one year.

## Missing data due to death

- However, some patients in the trial died before the one year since the receipt of the treatment, and hence their health related quality of life measures are not well defined.
- An one-year death rate was 21.3% in the DE and 28% in MP regimen, respectively.

## Existing Methods - naive method

- One of the naive analytic approaches is to perform analysis only on subjects who are alive or on all observed data.
- A major limitation of this approach is that it focuses on the healthiest subgroup of population (the survivors) and typically induces bias in the inference on the original cohort.

## Existing Methods - missing-data methods

- Data truncated by death are sometimes handled by a method for missing data.
- Revicki et al. (2001) imputed the data that would have been observed if the subject had not died.
- However, there is philosophical distinction between outcomes truncated by death and outcomes missing due to dropout. If a subject dies, all subsequent outcomes no longer exist and are therefore considered undefined.
- Rubin (2000, 2006) has convincingly argued that the approach of treating undefined quality of life outcomes as missing data is inappropriate and misses the scientific, medical, and ethical value of the quality of life outcome.

## Existing Methods - transformation

- Another approach is to transform the outcome of interest into a new variable that has a clear meaning and is defined for the death group.
- Diehr et al. (2005) adopted this approach by transforming the eleven health-related variables into new variables to study the cardiovascular health for adults aged 65 or older.
- However, quantifying a value to quality of life of a patient who is dead can also be subjective and problematic.



# Three Survivor-based Approaches

- Fully conditional: modeling the conditional distribution of QOL on a given survival time and describing longitudinal change in QOL separately for each survival time point.
- Partly conditional: modeling conditional distribution of QOL on being alive at a time and describing the relationship between risk factors and QOL in survivors at a particular time.
- Principal stratification, conditioning on survivors.

# Different Interpretation

- Fully conditioning model can tell us individual trajectories, but uses future survival information to predict earlier responses
- Partly conditioning model can tell us the dynamic relationship between risk factors and QOL over dynamic cohort (Kurland, et al, 2009)
- Principal stratification is most suited for estimating treatment effects.

## Survivor average causal effect (SACE)

- Rubin (2000, 2006) argued that a better approach would be using the potential outcome framework to estimate a clinically meaningful causal effect parameter, called "survivor average causal effect" (SACE), among patients who would survive under the both treatments.

## Existing Methods, continued

- Zhang et al. (2009) construct a parametric mixture normal model and presented the maximize likelihood estimation.
- A big issue with truncation by death is model identifiability.
- Even the parametric mixture normal model is not identifiable, if the probability of each latent component is the same.
- When the outcome is binary, even if the probabilities of latent components are different, we cannot get identifiability for a binary mixture model.

# New semi-parametric and non-parametric Methods

- In this talk, we focus on the identifiability of the principal causal effect of interest in a non-parametric or semi-parametric model and propose some conditions and assumptions which ensure identifiability.
- Estimation.

# Notations

- Let  $Z$  denote the randomized treatment assignment (1 for a new treatment, 0 for a control treatment), and let  $S(z)$  be the potential survival indicator (1 for survival, 0 for death) if the patient is assigned to treatment  $z$ .
- Let  $Y(z)$  denote a measurement of the potential quality of life if a patient is assigned to treatment  $z$ .
- Foundational problem of causal inference: only  $(Y(1), S(1))$  or  $(Y(0), S(0))$  is observed.
- Let  $S = S(Z)$  and  $Y = Y(Z)$  denote the observed survival status and the observed QOL, respectively.

# Principal Strata

- Let  $G$  be the principal strata of patients, which is defined as follows:

$$G = \begin{cases} LL, & \text{if } S(1) = 1 \text{ and } S(0) = 1; \\ LD, & \text{if } S(1) = 1 \text{ and } S(0) = 0; \\ DL, & \text{if } S(1) = 0 \text{ and } S(0) = 1; \\ DD, & \text{if } S(1) = 0 \text{ and } S(0) = 0. \end{cases}$$

Define  $\pi_g = P(G = g)$  for  $g = LL, LD, DL$  or  $DD$ .

- Causal parameter within each stratum:

$$ACE_g = E(Y(1) - Y(0) \mid G = g),$$

where  $g = LL, LD, DL, DD$ .

## Parameter of interest

- For other principal strata  $g = LD, DL$  and  $DD$ , we cannot define a meaningful  $ACE_g$  because there exists no real valued QOL either for control group ( $Z = 0$ ) or treatment group ( $Z = 1$ ).
- Thus we focus only on the meaningful  $ACE_{LL}$ .
- We consider the average causal effect  $ACE_{LL}$  for the principal stratum  $LL$ , (Rubin (2006) call it  $SACE$ ), as the parameter of interest, defined as  $ACE_{LL} = E\{Y(1) - Y(0) \mid G = LL\}$ .



# Assumptions

- **Assumption 1:** Stable unit treatment value assumption (SUTVA). There is no interference between units, which means that the potential outcomes of one individual do not depend on the treatment status of other individuals (Rubin, 1980).
- **Assumption 2** (Randomization):  $Z$  is randomized.
- **Assumption 3** (Monotonicity):  $S_i(1) \geq S_i(0)$  for all subjects. In our example,  $1 = DE$  and  $0 = MP$ .

## Identifying causal effects with help of a covariate

- To identify the principal causal effect  $ACE_{LL}$ , we introduce a covariate  $A$  which is pretreatment and satisfies the following assumption 4.
- **Assumption 4:** The pre-treatment covariate,  $A$ , is independent of  $Y$  given  $G$  and  $Z$ , denoted as  $A \perp Y | (Z, G)$ .
- Assumption 4 implies 
$$E\{Y | Z = 1, G, A\} = E\{Y | Z = 1, G\}.$$
- The pre-treatment covariate,  $A$ , defined in this assumption, is similar to but different from an instrumental variable.

## Identifiability under assumptions 1-5

- **Theorem 1.** Under Assumptions 1 to 4, the principal causal effect  $ACE_{LL}$  is identifiable if the covariate  $A$  is predictive to the strata  $LL$  and  $LD$ , that is,  
 $P(A = a \mid G = LL) \neq P(A = a \mid G = LD)$  for some  $a$ .

# Moment estimation under assumptions 1-5

- We propose a moment estimation method for estimating  $ACE_{LL}$  under Assumptions 1 to 5.

# Application to SWOG

- In our analysis,  $Z = 1$  denotes DE and  $Z = 0$  denotes MP. We use the baseline QOL and the difference between QOL after one year and the baseline QOL as  $A$  and  $Y$ , respectively.
- $S = 0$  if a patient died within one year after the receipt of a treatment; and otherwise  $S = 1$ .

# Application to SWOG

- We analyze the data using the three different methods, corresponding to the following three cases with or without Assumptions 3 and 4, using the baseline QOL score as the covariate  $A$ :
  1. The proposed method under the Assumptions 1 to 4.
  2. The proposed method without the monotonicity assumption 3. For the continuous  $A$ , we discretise it into a three-level variable  $A'$  in two ways: one way is  $A' = 1$  for  $A \leq 25$ ,  $A' = 2$  for  $25 < A \leq 75$  and  $A' = 3$  for  $A \geq 75$ ; the other way is  $A' = 1$  for  $A \leq 50$ ,  $A' = 2$  for  $50 < A \leq 70$  and  $A' = 3$  for  $A \geq 75$ .
  3. The proposed method without the independence assumption 4, using the linear model.

# Result

Case	$\widehat{ACE}_{LL}$	s.e.	2.5%	50%	97.5%
1	7.01	3.09	1.81	6.56	13.64
2 (for discretization 2)	-2.20	6.07	-12.27	0.08	12.45
2 (for discretization 1)	0.63	5.45	-11.34	0.06	10.23
3	6.97	4.30	-4.68	2.96	12.49

## Application to SWOG, continued

- From Table, we can see that cases 1 and 3 get quite different results from case 2.
- This result suggests that the estimates are sensitive to the monotonicity assumption.
- In case 2, we get the estimate of  $\pi_{DL}$  greater than 0.33. This may be an evidence for the violation of the monotonicity assumption since the estimate of  $\pi_{DL}$  usually should be very small if the monotonicity assumption holds.



## Application to SWOG, continued

- Our results show that the monotonicity assumption in this example plays an important role.
- Under this assumption, we may conclude the treatment has a significant causal effect on the quality of life among the LL group.
- Without this assumption, we will not be able to reach this positive conclusion.

# Acknowledgement

- This is a joint work with Zhi Geng, Ding Peng, Wei Yuan at Peking University.
- This work has been published in Journal of American Statistical Association.