How to analyze the duration of response in oncology trial?

Lu Tian

Stanford University

Joint Work with Bo Huang (Pfizer), and LJ Wei (Harvard)

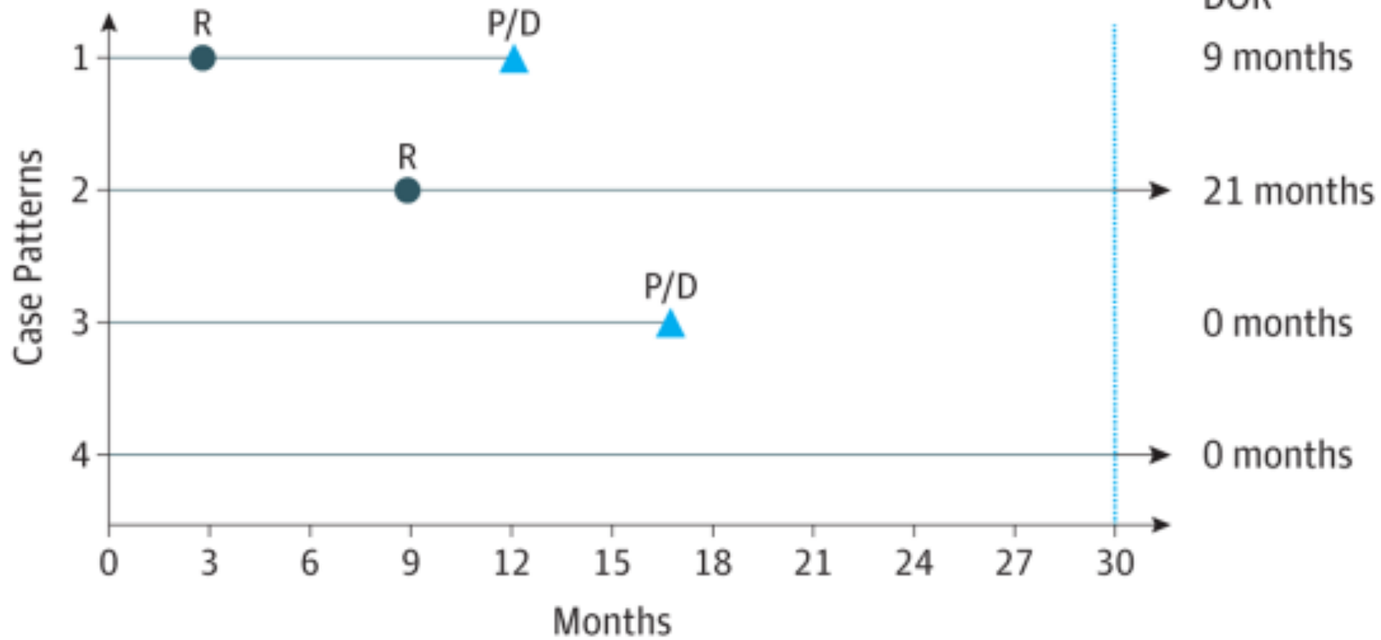
15th Annual Conference on Statistical Issues in Clinical Trials

April 17, 2023

Endpoints

- Time to response (R)
- Time to progression or death (P/D)
- Duration of Response (DOR)
  - No response after progression or death, i.e., P/D is an absorbing state
  - P/D may occur first, and time to response can be defined as infinity in such a case.

Duration of Response



RMST Analysis for Duration of Response

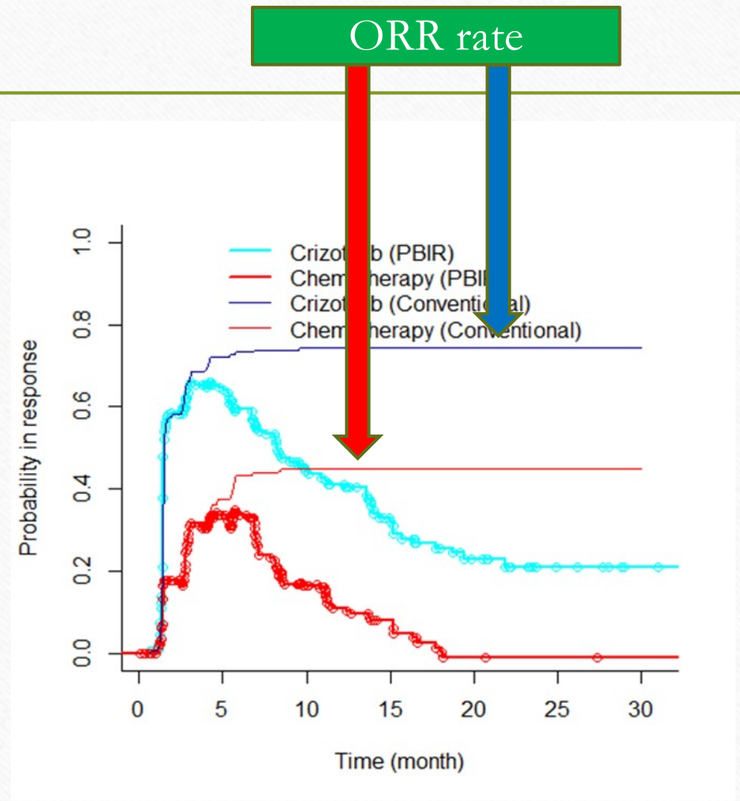
- Conventional Approach
  - We often focus on the cumulative overall response rate (ORR) in Phase 2 studies.
  - We estimate the distribution of the duration of response using KM estimator among responders
- The approach fails in several aspects:
  - Post-response progression isn't considered in reporting ORR
  - The KM estimator is biased because of informative censoring
  - Not clear how to combine ORR and KM estimator among responders to assist informative clinical decision (effective or not)

Example: PROFILE-1014

- A multicenter, randomized, open-label, phase 3 study that compared Crizotinib and pemetrexed-plus-platinum chemotherapy with respect to efficacy and safety
- 343 patients with previously untreated, advanced, ALK-positive non-small cell lung cancer (NSCLC).
- Patients were randomly assigned, in a 1:1 ratio, to receive oral crizotinib (N=172), or intravenous chemotherapy (N=171, pemetrexed plus either cisplatin or carboplatin).
- The primary endpoint was progression-free survival (PFS).
- The time to OR and the DOR were major secondary endpoints.
- All efficacy analyses were performed using data collected by Month 30.

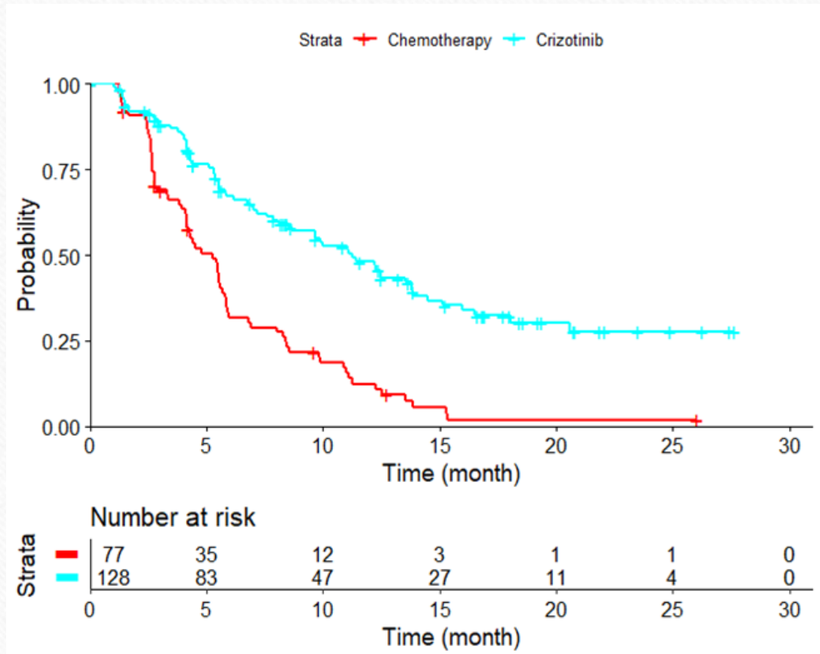
Conventional Analysis

- The cumulative ORRs at Month 6 were 73% (=126/172) and 43% (=74/171), for Crizotinib and chemotherapy, respectively.
  - Treating all censored events as non-responders
- The cumulative ORR curves in Figure 2 plateau after 6 months, suggesting that almost all of the ORs occurred before Month 6.
- The ORR curves always increase even though many patients stop responding to the treatment after initial response.



Conventional Approach

- KM curves of the DOR among responders.
- The median DORs were 11.3 months for Crizotinib and 5.3 months for chemotherapy.
- No formal analysis would be conducted to compare the two KM curves.



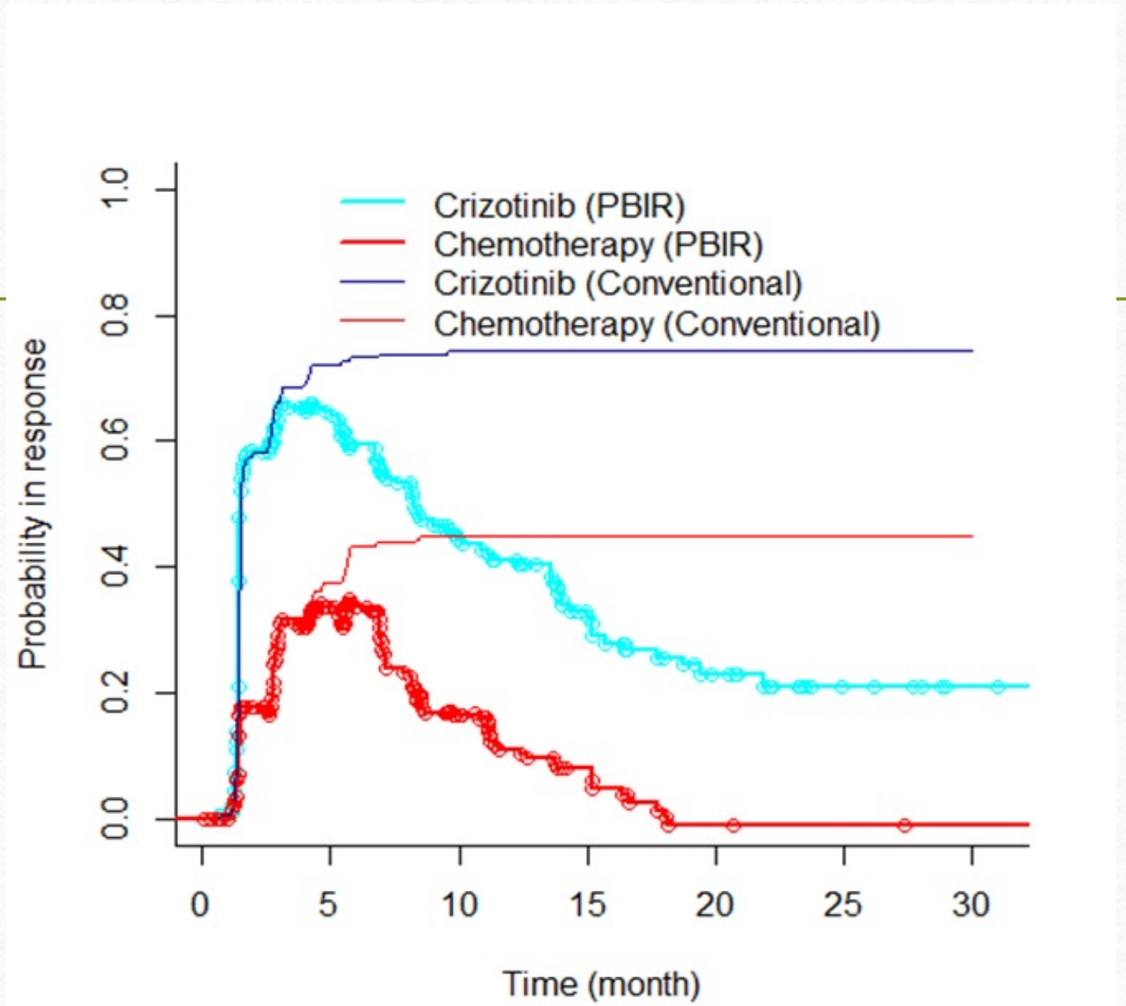
Why is the KM Estimator biased?

- T_R: time to response from randomization
- T_D: time to progression from randomization
- C: time to censoring (drop out or study termination)
- The KM estimator is based on the subgroup of patients with $T_R < T_D$
  - $X = \min(T_D - T_R, C - T_R)$
  - $\delta_X = 1(T_D - T_R < C - T_R) = 1(T_D < T_R)$
  - The objective is to estimate $S_{D|R}(t) = P(T_D - T_R > t | T_D - T_R > 0)$
  - However, $S_{D|R}(\cdot)$ can not be estimated via a simple KM estimator since
    - $T_D \perp C \Rightarrow T_D - T_R$ and $C - T_R$ are dependent

- A more meaningful endpoint is
  - **the time length a responder staying as a responder (DOR)**
  - **the proportion of responders at a particular time point**

THE PROBABILITY OF BEING IN RESPONSE (PBIR)

- The occurrence of ORs competes with the occurrence of P/D.
- In the absence of censoring, both estimation of the PBIR at any given time point and construction of a confidence interval for the difference in PBIRs are straightforward:
  - $PBIR(t) = P(T_D > t, T_R < t) = P(T_D > t) - P(T_D \wedge T_R > t)$
  - KM estimators for T_D and $T_D \wedge T_R$: $\hat{S}_D(t), \hat{S}_{D \wedge R}(t)$
  - $\widehat{PBIR}\,(t) = \hat{S}_D(t) - \hat{S}_{DR}(t)$
  - $\widehat{PBIR}\,(t) - PBIR(t) \approx N(0, \sigma(t)^2)$
    - The variance needs to account for the correlation between two dependent KM estimators.


Crizotinib (PBIR)
Chemotherapy (PBIR)
Crizotinib (Conventional)
Chemotherapy (Conventional)
Probability in response
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
30
Time (month)

THE PROBABILITY OF BEING IN RESPONSE (PBIR)

- Month 6
  - the PBIRs were 60% and 34% for Crizotinib and chemotherapy, respectively.
  - The difference (Crizotinib minus chemotherapy) was 26% in favor of Crizotinib, with 95% confidence interval (CI) 17% to 34%.
- Month 12,
  - the PBIRs were 41% and 11%.
  - The difference was 30% in favor of Crizotinib, with 95% CI 22% to 38%.
- Over the entire follow-up period, the PBIR curve for Crizotinib is uniformly higher than that for chemotherapy.
- The peaks of the curves occurred at 4.2 months for Crizotinib and 4.6 months for chemotherapy. Unlike the cumulative ORR curves, the PBIR curves decline after the peaks, reflecting the fact that some patients did not have a sustained response. It is interesting to note that the rates of the decline were similar for two study therapies.

Duration of Response (DOR)

- Assumption: no response after progression
  - T_D: Time to progression or death (progression free survival)
  - T_R: Time to response.
- Duration of response: $(T_D - T_R)I(T_R < T_D) = T_D - T_D \wedge T_R$
- Restricted Duration of Response (RDOR): $\mathrm{D}_\tau = T_D \wedge \tau - T_D \wedge T_R \wedge \tau$
- Restricted Mean Duration of Response (RMDOR): Consider the average response time within a window $[0, \tau]$

$$RMDOR = E(T_D \wedge \tau - T_D \wedge T_R \wedge \tau) = \int_0^\tau PBIR(t)dt$$

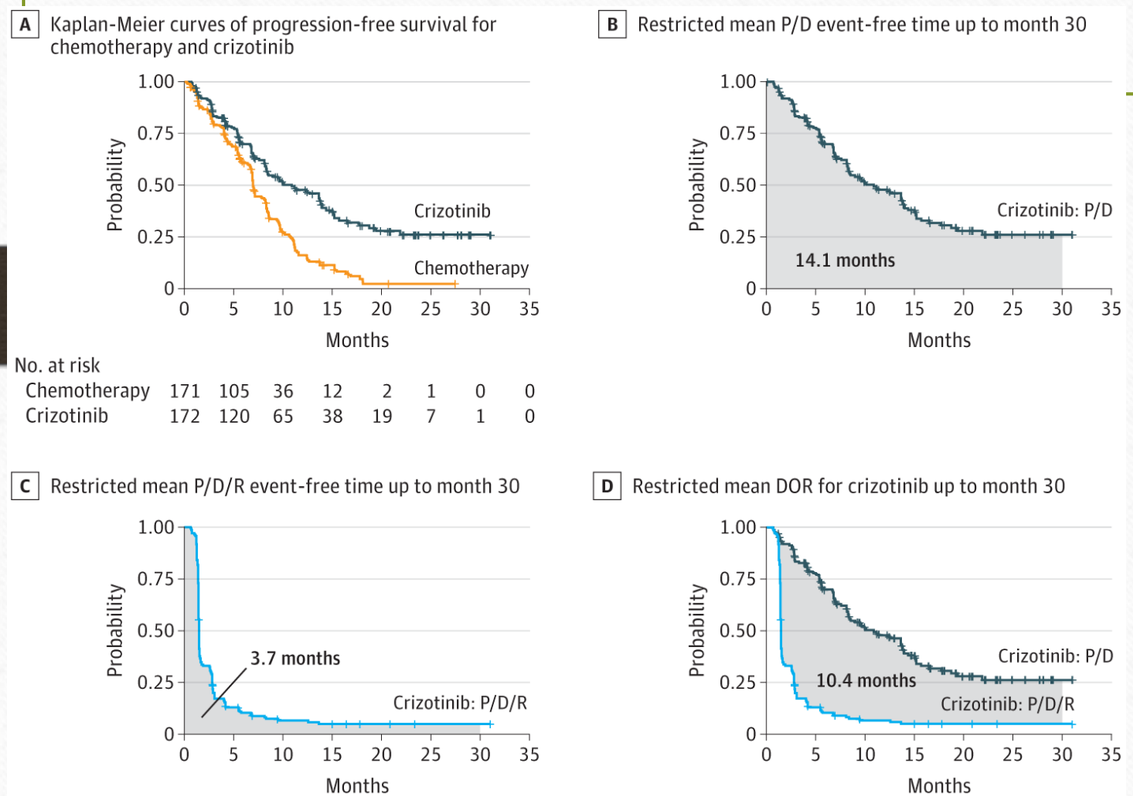
Inference for the RMDOR

- A consistent estimator for RMDOR:

$$\int_0^\tau \widehat{PBIR}(t)dt = \int_0^\tau \{\hat{S}_D(t) - \hat{S}_{DR}(t)\}dt,$$

  - $\hat{S}_D(\cdot)$ is the KM estimator for the survival function of T_D
  - $\hat{S}_{DR}(\cdot)$ is the KM estimator for the survival function of time $T_D \wedge T_R$
- The restricted mean duration of response can be represented by the area between two KM curves.

DMDOR Example



- crizotinib vs chemotherapy for patients with ALK positive lung cancer
- PFS of crizotinib is longer than that with chemotherapy (HR=0.45, 95% CI 0.350.60; p<0.001)
- Objective RR for crizotinib and chemo are 74% and 45%, respectively.
- Consider the restricted mean DOR within [0, 30] months:
  - On average patients with crizotinib have 10.4 months DOR over a 30 month follow-up
  - On average patients with chemotherapy have 3 months DOR over a 30 months follow-up
  - The difference in DOR is 7.4 months (95% CI 6.0-8.8 months, p<0.001)

How about the survival distribution of the DOR?

- One may be interested in estimating the survival probability that

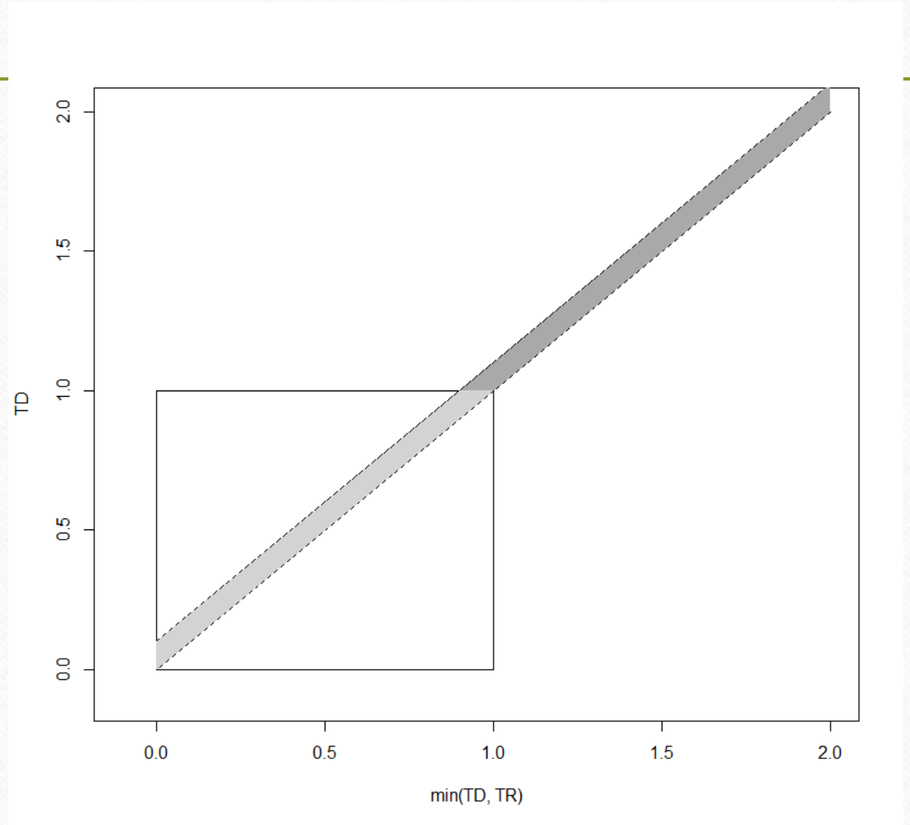
$$P(T_D - T_R > t | T_D > T_R)$$

or

$$P(T_D - T_R \wedge T_D > t)$$

- Both neither was identifiable in general.
  - The best we can know is the joint distribution of $(T_R, T_R \wedge T_D)'$ within a region $[0, \tau] \times [0, \tau]$, from which it is IMPOSSIBLE to derive the distribution of $(T_D - T_D \wedge T_R)$.

Non-identifiability



Survival Function of RDOR

- The survival distribution of $D_\tau = T_D \wedge \tau - T_R \wedge T_D \wedge \tau$ (RDOR within $[0, \tau]$)

$$S_D(t) = P(D_\tau > t)$$

  - If the support of T_R is $[0, \tau_R] \subset [0, \tau]$, then for $t \in [0, \tau - \tau_R]$

$$S_D(t) = P(T_D - T_R \wedge T_D > t)$$

  - The survival distribution of restricted DOR among responders

$$S_D(t|0) = P(D_\tau > t | D_\tau > 0) = S_D(t)/S_D(0)$$

- There is a class of estimators for estimating $S_D(t)$.

A Simple IPW Estimator

- Observed data: $(X_{1i}, X_{2i}, \delta_{1i}, \delta_{2i}), i = 1, \cdots, n$
  - $X_{2i} = T_{Di} \wedge \tau \wedge C_i$
  - $\delta_{2i} = I(T_{Di} \wedge \tau < C_i)$
  - $X_{1i} = T_{Di} \wedge T_{Ri} \wedge \tau \wedge C_i$
  - $\delta_{1i} = I(T_{Di} \wedge T_{Ri} \wedge \tau < C_i)$
  - $D_i = T_{Di} \wedge \tau - T_{Di} \wedge T_{Ri} \wedge \tau$

- The initial IPW estimator

$$\hat{S}_1(t) = n^{-1} \sum_{i=1}^{n} \frac{\delta_{2i}}{\hat{G}_C(T_{2i} \wedge \tau)} I(D_i > t),$$

where $\hat{G}_C(\cdot)$ is the KM estimator for the distribution of C_i.

An Improved IPW Estimator

- Key observation: $\delta_{2i} = 1 \Rightarrow D_i$ is observable $\Rightarrow I(D_i > t)$ is observable.
- But we don't need to know the precise value of D_i in order to determine if $(D_i > t)$
  - $\delta_{2i} = 0, \delta_{1i} = 1, X_{2i} > X_{1i} + t \Leftrightarrow T_{Di} > C_i > T_{Ri} + t \Rightarrow D_i > t$
- The improved IPW estimator:

$$\hat{S}_2(t) = n^{-1} \sum_{i=1}^{n} \frac{I(C_i > T_{Di} \wedge (T_{Ri} + t) \wedge \tau)}{\hat{G}_C(T_{Di} \wedge (T_{Ri} + t) \wedge \tau)} I(D_i > t)$$

Improved IPW Estimator

- Key Observation:

$$\frac{I(C_i > T_{Di} \wedge (T_{Ri} + t) \wedge \tau)}{\hat{G}_C(T_{2i} \wedge (T_{Ri} + t) \wedge \tau)} I(D_i > t) = \frac{(\delta_{2i} + (1 - \delta_{2i})\delta_{1i}) I(X_{2i} - X_{1i} > t)}{\hat{G}_C(X_{2i} \wedge (X_{1i} + t) \wedge \tau)}$$

  is always observable.

- Since $I(C_i > T_{Di} \wedge (T_{Ri} + t) \wedge \tau)$ has more nonzero terms than δ_{2i}, the new IPW estimator is expected to be more efficient

Connection with Lin and Ying estimator

- Lin and Ying has proposed an estimator for the distribution of gap time in recurrent events setting:

$$\hat{S}_{LY}(t) = n^{-1} \sum_{i=1}^{n} \frac{I(X_{2i} - X_{1i} > t)}{\hat{G}_C(X_{1i} + t)}$$

- Rational:

$$E\left(\frac{I(X_{2i} > X_{1i} + t)}{\hat{G}_C(X_{1i} + t)}\right)$$

$$\approx P(T_{Di} \wedge \tau > T_{Ri} \wedge T_{Di} \wedge \tau + t) \frac{P(C_i > T_{Ri} \wedge T_{Di} \wedge \tau + t)}{G_C(T_{Ri} \wedge T_{Di} \wedge \tau + t)} = P(D_i > t)$$

- It turns out that $\hat{S}_{LY}(t)$ and the improved IPW estimator $\hat{S}_2(t)$ are identical!

Nonparametric Regression

- Observing the fact that

$$S_{D|R}(t) = \int_0^\tau S_D(t|s)dF_{DR}(s),$$

where

$F_{DR}(s) = P(T_R \wedge T_D \wedge \tau \leq s) = 1 - S_{DR}(s)$;
$S_{D|R}(t|s) = P(D_\tau > t|T_R \wedge T_D \wedge \tau = s)$.

- We may estimate $S_D(t)$ by

$$\hat{S}_3(t) = \int_0^\tau \hat{S}_{D|R}(t|s)d\{1 - \hat{S}_{DR}(s)\},$$

  - $\hat{S}_{DR}(s)$ is the regular KM estimator for the survival function of $T_R \wedge T_D \wedge \tau$
  - $\hat{S}_D(t|s)$ is a nonparametric consistent estimator for $S_D(t|s)$

The Construction of $\hat{S}_D(t|s)$

- Noting that

$$P(D > t|T_R \wedge T_D \wedge \tau = s)$$
$$= P(D > t|T_R \wedge T_D \wedge \tau = s,\ C > T_R \wedge T_D \wedge \tau)$$
$$= P(D > t|X_1 = s, \delta_1 = 1),$$

we can estimate $S_{D|R}(t|s)$ by a kernel smoothed KM estimator:

$$\hat{S}_{D|R}(t|s) = \exp\left\{-\int_0^t \frac{\delta_{2i} d \sum_{i=1}^n I(D_i \le u)\delta_{1i}K_h(X_{1i} - s)}{\sum_{i=1}^n I(X_{2i} - X_{1i} \ge u)\delta_{1i}K_h(X_{1i} - s)}\right\}.$$

Summary

- A simple IPW estimator $\hat{S}_1(t)$
- An improved IPW estimator $\hat{S}_2(t) = \hat{S}_{LY}(t)$
- A regression estimator $\hat{S}_3(t)$
- All those estimators are consistent and asymptotically normal with the first order approximation:

$$\sqrt{n}\{\hat{S}_j(t) - S_D(t)\} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_{ij}(t) + o_p(1), j = 1,2,3.$$

- The variance of $\hat{S}_j(t)$ can be estimated by

$$\frac{1}{n^2}\sum_{i=1}^{n}\hat{\xi}_{ij}(t)^2, j = 1,2,3.$$

Connection with RDOR

- Recall that RMDOR can be estimated by

$$\hat{\mu}_D = \int_0^\tau \{\hat{S}_D(t) - \hat{S}_{DR}(t)\} = \int_0^\tau \widehat{PBIR}(t)dt\,.$$

- On the other hand, the RMDOR also can be estimated by

$$\int_0^\tau \hat{S}_j(t)dt\,, j = 1, 2, 3.$$

- If an estimator for the survival function of D_τ is "good", then the corresponding estimator for RMDOR should be "good" as well. What if it is not "good"?

Efficiency Augmentation

- Consider the statistic:

$$\hat{\eta} = \int_0^\tau \hat{S}_1(t)dt - \left[\int_0^\tau \widehat{PBIR}(t)dt\right]$$

we have

$$\sqrt{n}\begin{pmatrix}\hat{S}_1(t) - S_D(t)\\ \hat{\eta}\end{pmatrix} \approx \frac{1}{\sqrt{n}}\sum_{i=1}^n \begin{pmatrix}\xi_{1i}(t)\\ \tau_i\end{pmatrix} + o_p(1)$$

$$\rightarrow N\left(0, \begin{pmatrix}\sigma_{11}^2(t) & \rho(t)\sigma_{11}(t)\sigma_{22}\\ \rho(t)\sigma_{11}(t)\sigma_{22} & \sigma_{22}^2\end{pmatrix}\right)$$

Efficiency Augmentation

- Consider an augmented estimator

$$\hat{S}_{aug}(t) = \hat{S}_1(t) - \frac{\rho(t)\sigma_{11}(t)}{\sigma_{22}}\hat{\eta} \sim N\left(S_D(t), \sigma_{11}^2(t)\left(1-\rho(t)\right)^2\right)$$

  which is more efficient than the initial estimator $\hat{S}_1(\cdot)$ and consistent.

- In practice, one needs to estimate $\rho(t)$, $\sigma_{11}(t)$, and σ_{22}, and the augmented estimator can be written as

$$\hat{S}_1(t) - \frac{\hat{\rho}(t)\hat{\sigma}_{11}(t)}{\hat{\sigma}_{22}}\hat{\eta} = \hat{S}_1(t) - \hat{w}(t)\hat{\eta}$$

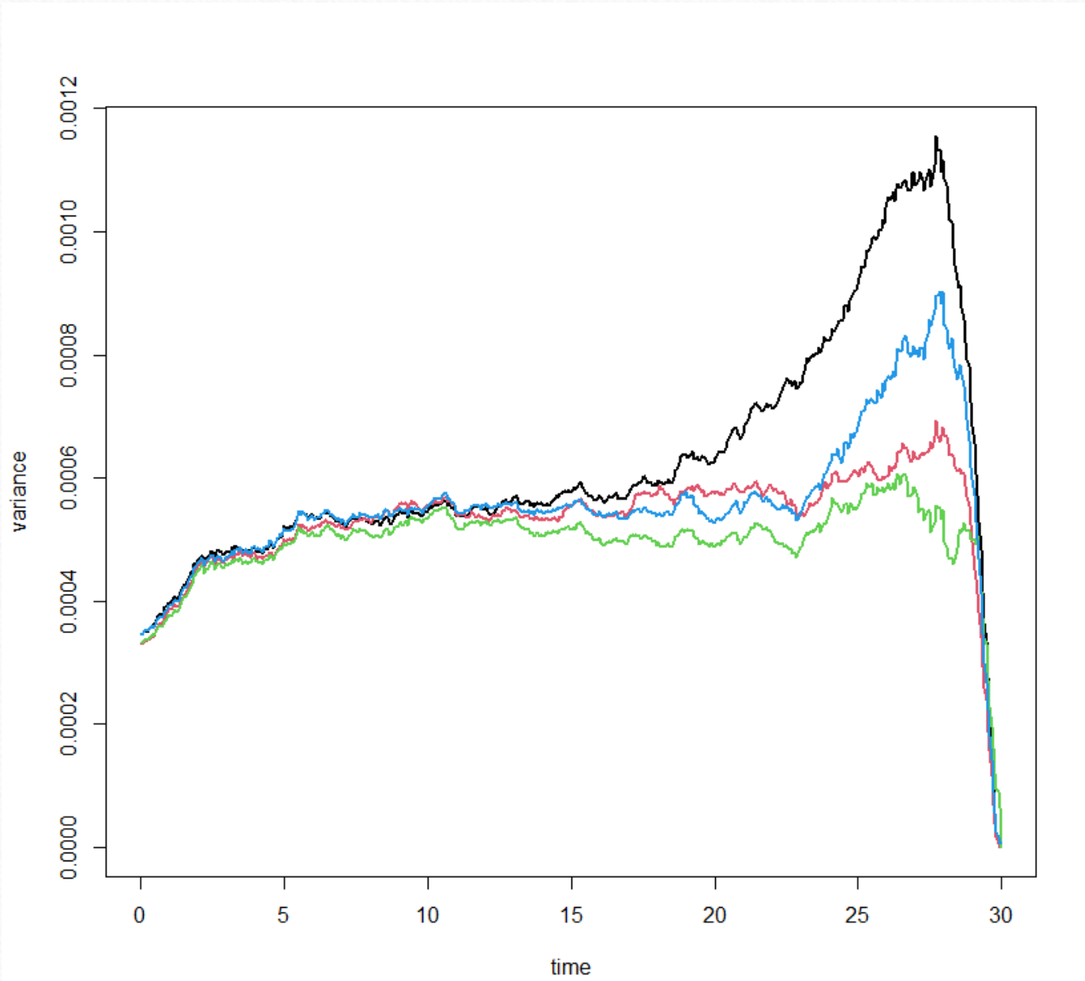
Empirical variance of four estimators

IPW1

IPW2

SMOOTHED (BEST)

IPW1+Augmentation

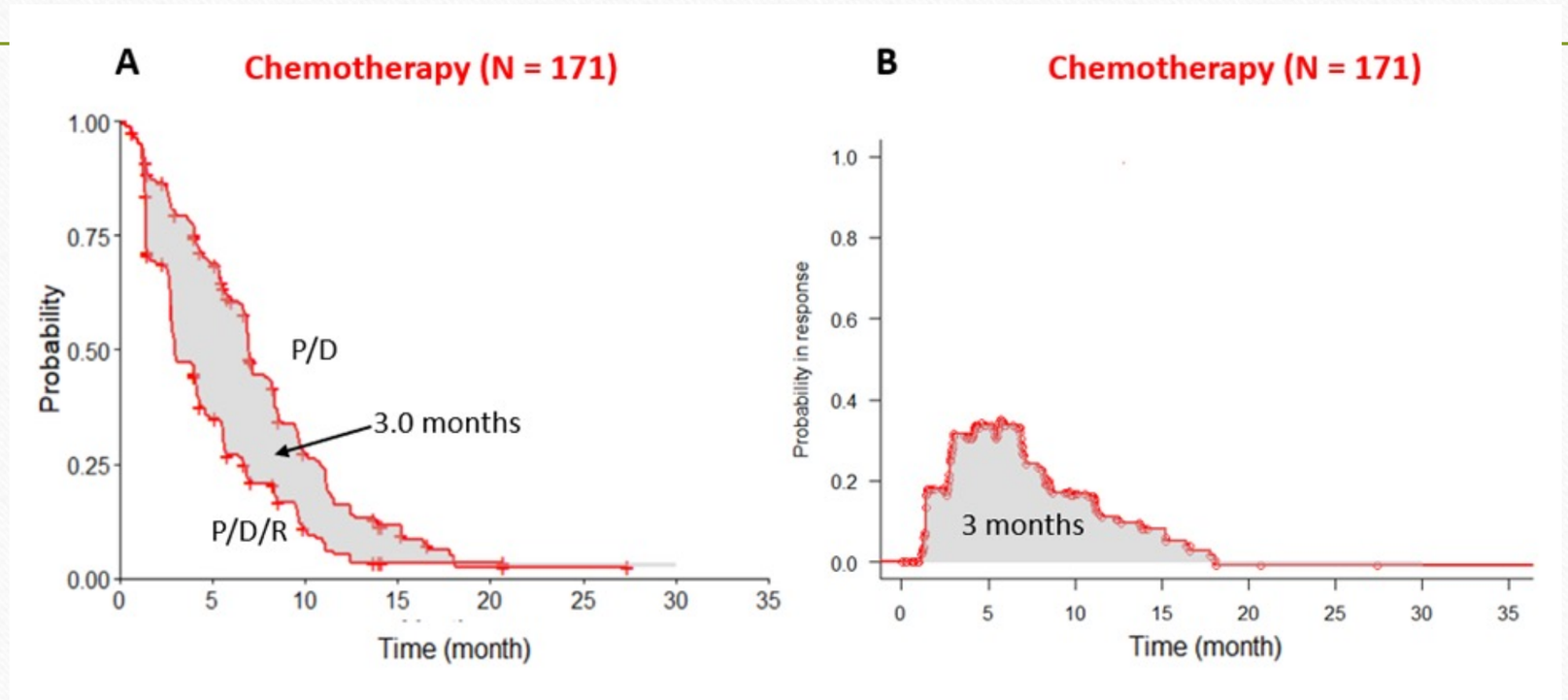


Summary

- Most analytical procedures discussed in this paper can be implemented via publicly available software: https://web.stanford.edu/~lutian/Software.HTML.
- We should perform more "correct" and less "ad-hoc" analysis.
- There may be more efficient estimators for $PBIR(t)$, RMDOR and the survival distribution of RDOR.

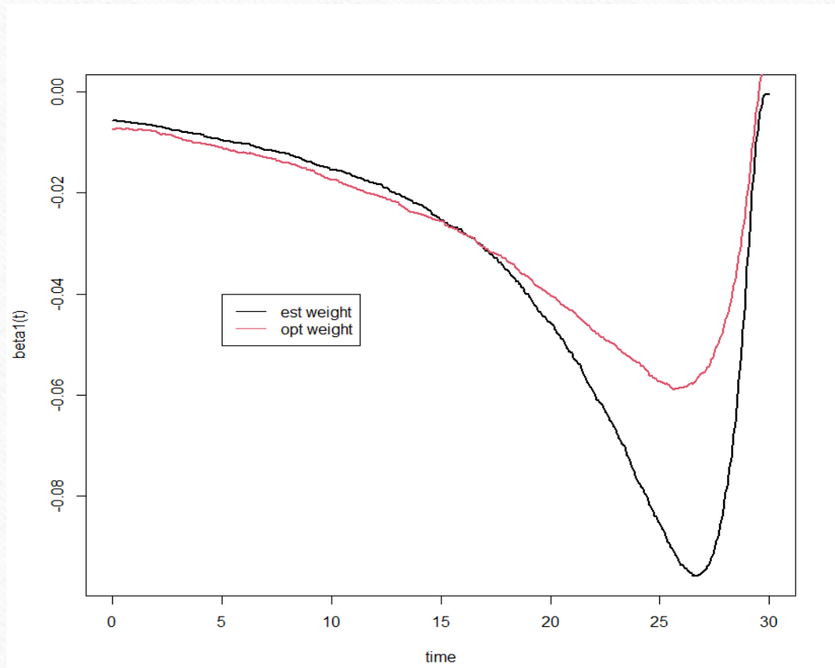
BACK UP SLIDES

A: 30-month mDOR as the area between two Kaplan-Meier curves, i.e., the area between the progression/death (P/D) curve and the progression/death/response (P/D/R) curve.
B: 30-month mDOR as the area under PBIR curve



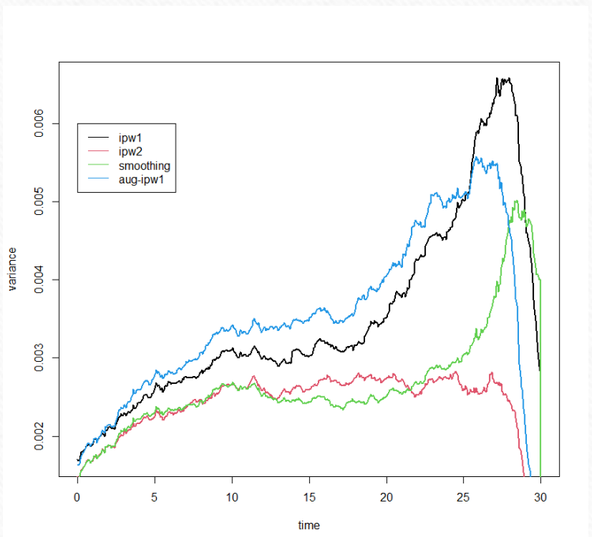
Efficiency gain is not free

- The bias and variability in estimating the optimal $\hat{\beta}_1(t)$ may destroy the efficiency gain

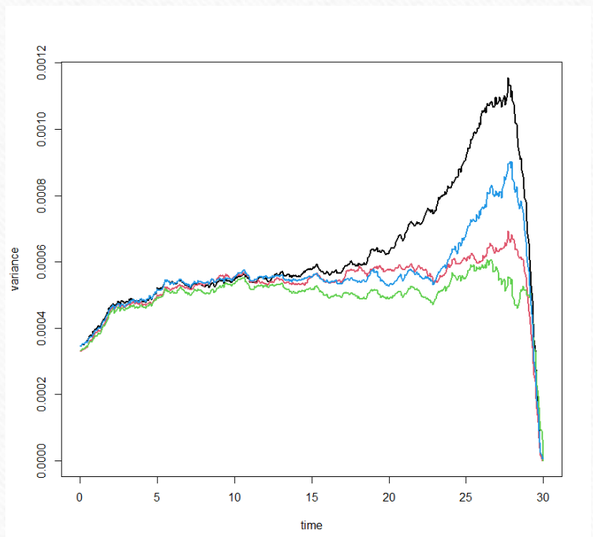


The price of augmentation

Sample size of 100

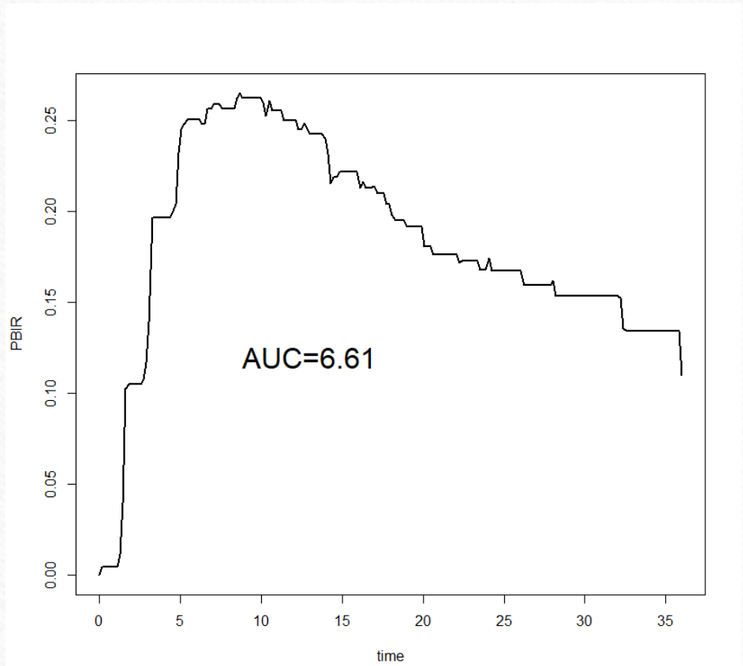


Sample size of 500



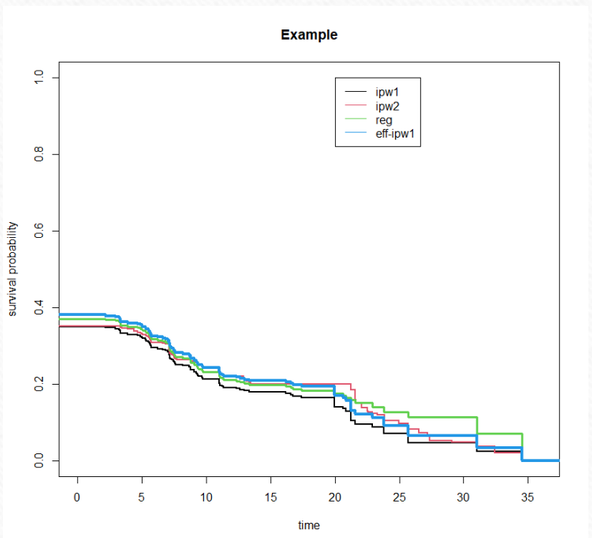
Example

- Simulated data set: 444 patients, 138 responses and 284 progressions
- PBIR Curve
- Estimated RMDOR: 6.61 (95% CI: 5.51-7.82) months.

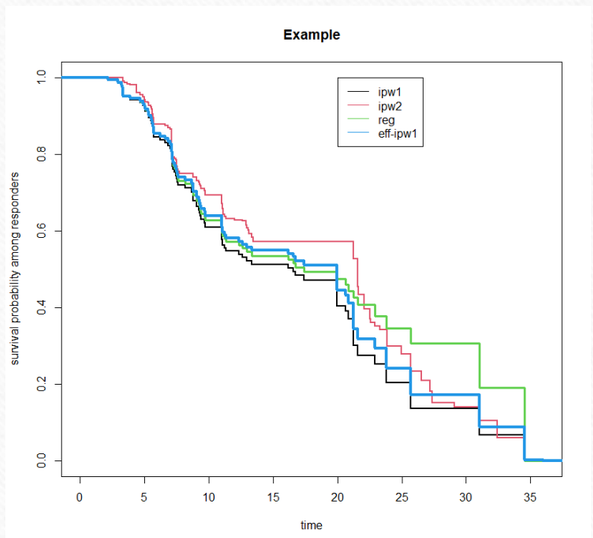


Example

Survival function for RDOR

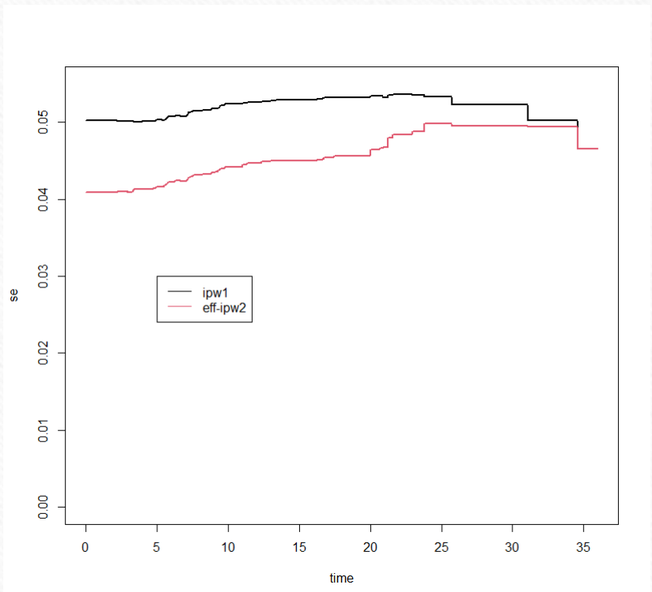


Survival function for RDOR among responders



Effect of efficiency augmentation

Standard Error of Two Estimators



- There are substantial efficiency gains at some time points.
- RMDOR estimate based on the naïve estimator is 5.74 months
- RMDOR estimate based on the PBIR is 6.61 months
- RMDOR estimate based on the efficiency-augmented estimator is 6.61 months

A Simulation Study

- Distribution Assumptions
  - $(T_1^*, T_2^*) \sim e^{N(\mu,\Sigma)}$
    - $\mu = (1.2, 2)'$
    - $\Sigma = 2\begin{pmatrix} 1.0 & 0.3 \\ 0.3 & 1.0 \end{pmatrix}$
  - $(T_R, T_D) = (T_1^* \wedge T_2^* ,\ T_2^*)$
  - $C = U(12, 24) \wedge EXP\left(\frac{1}{300}\right)$
- KM curve $(n = 10^6)$ vs. $P(T_D - T_R > t | T_D > T_R)$
- The bias can be nontrivial for the second half of the survival curve.