Interim Analysis of SMART Survival Data

Zi Wang

University of Pittsburgh

Mar 24, 2025

Disclaimer

- Research reported in this presentation was partially funded through a Patient-Centered Outcomes Research Institute (PCORI) Award (ME-2021C3-24215).
- The statements in this presentation are solely the responsibility of the authors and do not necessarily represent the views of PCORI, its Board of Governors or the Methodology Committee.

Background and Motivation

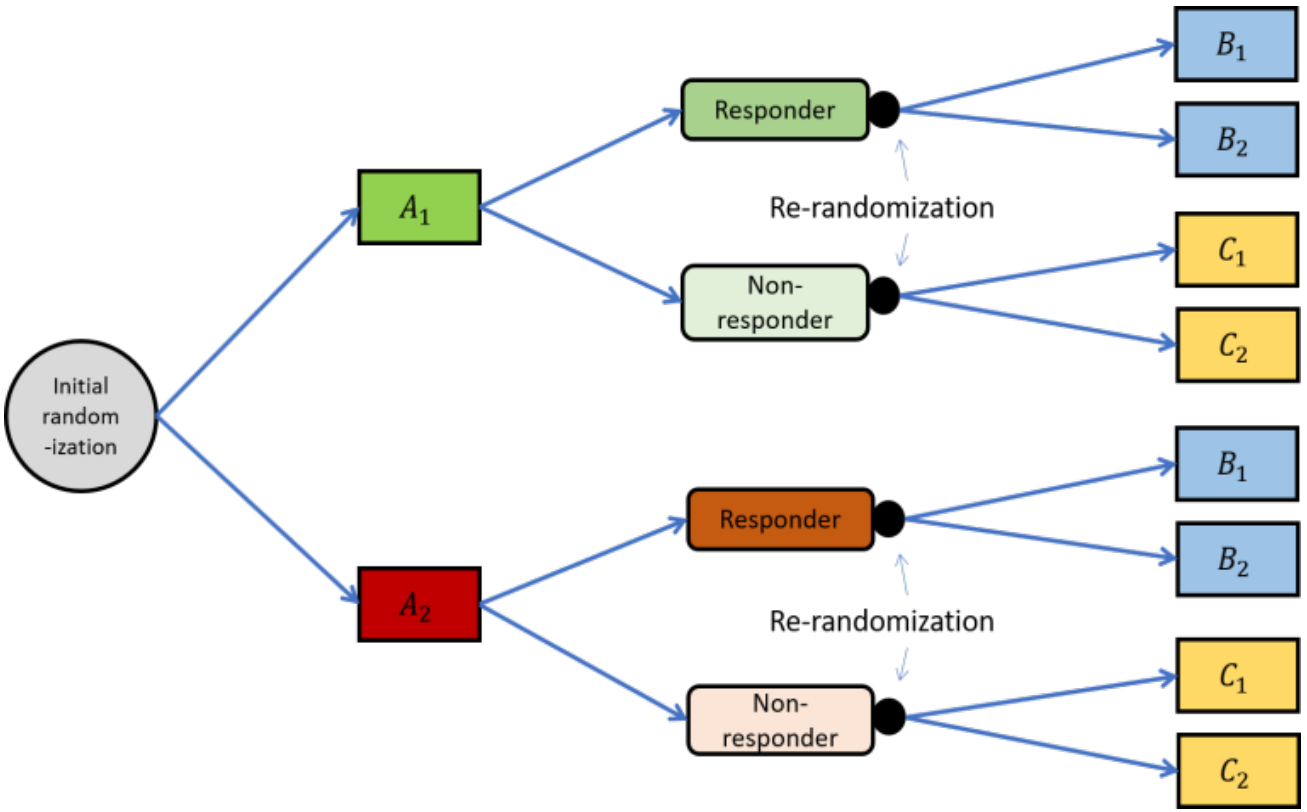
- A dynamic treatment regime (DTR) is a mathematical representation of a multistage decision process.
- Sequential multiple assignment randomized trial (SMART) designs provide a systematic process of constructing and evaluating DTRs.
- In a SMART, patients are involved in multiple stages of treatment, and the treatment assignment is adapted over time based on the patient's characteristics and treatment history.

Interim Monitoring (IM)

- To improve the efficiency, consider interim monitoring (IM)
  - IM allows early stopping for superiority, inferiority, or futility.
- IM in SMARTs for testing multiple DTRs has received less attention.
- Some IM approaches were proposed for SMARTs.
  - Wu et al. (2023) proposed IM in SMART for continuous outcomes using patients who have completed all treatment stages.
  - Manschot et al. (2023) handles IM for SMART continuous data by using partial information from all patients.
- In this work, we develop and evaluate IM approaches to allow for early termination of SMART trials for survival outcomes.

Analysis for a 2-stage SMART with a Survival Outcome

- The treatment strategy $A_jB_kC_l$ is a decision rule prescribing how to treat a patient.



- Goal: compare the effects among treatment strategies $A_jB_kC_l$.

SMART Data

- Observed data:
  $\{I_i(A_j), \eta_i, \eta_i R_i, \eta_i R_i I_i(B_k), \eta_i(1-R_i)I_i(C_l), U_i, \delta_i\}, i = 1, \ldots, n$
  - η_i: the indicator for entering stage 2.
  - R_i: response indicator.
  - The observed time $U_i = \min(T_i, V_i)$, T_i the survival time and V_i the censoring time
  - The censoring indicator $\delta_i = I(T_i \leq V_i)$
- Observed event process: $N_i(s) = I\left(U_i \leq s, \delta_i = 1\right)$
- Observed at risk process: $Y_i(s) = I(U_i \geq s)$
- Total number of patients with an event at or before time s
  $N(s) = \sum_{i=1}^{n} N_i(s)$ and total number of patients at risk $Y(s)$

Interim Analysis for Survival Outcomes

- Interim analysis procedures have been well studied for survival outcomes.
  - Tsiatis (1982) developed the asymptotic joint distribution of test statistics from multiple interim analyses
  - Kim and Tsiatis (1990) developed IM procedure allowing for unequal increments between repeated analyses
  - Gu and Lai (1999): interim analyses based on Monte Carlo simulations.
  - Shen and Cai (2003): weighted average of the linear rank statistics.
- These methods cannot be directly applied to SMARTs.

Missingness by 2nd Stage Randomization

- If patients were assigned to $A_1B_1C_1$, all responders to A_1 would receive B_1, and all non-responders would receive C_1
- In SMART, responders were assigned to either B_1 or B_2, and non-responders were assigned to either C_1 or C_2.
- Inverse probability weighting (IPW) is needed to handle multi-stage randomization.

Inverse Probability Weighting (IPW)

- Consider time-dependent IPW for treatment strategy $A_jB_kC_l$:

$$W_{jkl,i}(s) = \frac{I_i(A_j)}{o_j}\left[1 - D_i(s) + D_i(s)\left\{\frac{R_i I_i(B_k)}{p_k} + \frac{(1-R_i)\,I_i(C_l)}{q_l}\right\}\right].$$

  - $D_i(s) = \eta_i I(T_{1i} < s)$, where T_{1i} is the time in stage 1.
  - o_j is the randomization probability of A_j, p_k is that for B_k, q_l is that for C_l.
- Define the weighted number of events by time s for $A_jB_kC_l$: $\bar{N}_{jkl}(s) = \sum_{i=1}^{n} W_{jkl,i}(s)N_i(s)$
- Similarly, we can define the weighted at-risk process $\bar{Y}_{jkl}(s)$.

Weighted Log-rank Statistics

- $\Lambda_{jkl}(t)$: Cumulative hazard w.r.t treatment strategy $A_jB_kC_l$.
  $\Lambda_0(t)$: Overall cumulative hazard under H_0.
- $H_0 : \Lambda_{jkl}(t)$ are all equal for $t \leq \tau$.
- χ^2 type test statistic for testing H_0: $T = n^{-1}\boldsymbol{Z}^\intercal \widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{Z}$.
  - $\boldsymbol{Z} = (Z_{jkl})^\intercal \in \mathbb{R}^7$, where the weighted log-rank statistic

$$Z_{jkl} = \int_0^\infty \frac{\bar{Y}_{jkl}(s)\bar{Y}_{111}(s)}{\bar{Y}_{jkl}(s) + \bar{Y}_{111}(s)} \left\{ \frac{\mathrm{d}\bar{N}_{jkl}(s)}{\bar{Y}_{jkl}(s)} - \frac{\mathrm{d}\bar{N}_{111}(s)}{\bar{Y}_{111}(s)} \right\}.$$

  - $\widehat{\boldsymbol{\Sigma}}$ is a consistent estimator of $\boldsymbol{\Sigma}$, the asymptotic covariance matrix of $n^{-1/2}\boldsymbol{Z}$.
- Under H_0, $T \sim \chi^2_\nu$ with df $\nu = rank(\boldsymbol{\Sigma}_0)$, where $\boldsymbol{\Sigma}_0$ is $\boldsymbol{\Sigma}$ evaluated under H_0.

Interim Monitoring in SMART

- M: total number of analyses, and t_m: decision time of the mth analysis.
- The test statistic at the mth analysis.

$$T(t_m) = n_m^{-1} \boldsymbol{Z}(t_m)^\intercal \hat{\boldsymbol{\Sigma}}^{-1}(t_m) \boldsymbol{Z}(t_m),$$

where $\boldsymbol{Z}(t)$ and $\hat{\boldsymbol{\Sigma}}(t)$ be $\boldsymbol{Z}$ and $\hat{\boldsymbol{\Sigma}}$ evaluated based on information up to time t.

- Efficacy boundaries $b_1, \ldots, b_M$ are determined to maintain the overall type I error, that is:

$$\sum_{m^*=1}^{M} \Pr\left(\bigcap_{m=1}^{m^*-1} T(t_m) \leq b_m \cap T(t_{m^*}) > b_{m^*} \mid H_0 \right) = \alpha.$$

  - Pocock boundaries (Pocock, 1977): $b_1 = \ldots = b_M$.
  - O'Brien Fleming (OBF) boundaries (O'Brien and Fleming, 1979): $b_m / b_M = \sqrt{M/m}$.

Interim Monitoring in SMART

Theorem (Joint Distribution of the Test Statistics)

Let $Q(t_m)$ *be a vector such that* $T(t_m) = Q(t_m)^\top Q(t_m)$ *for* $m = 1, \ldots, M$. *Under* H_0, $(Q(t_1), \ldots, Q(t_M))^\top \sim MN_{\nu \times M}(\mathbf{0}, \boldsymbol{\Psi})$, *where*

$$\boldsymbol{\Psi} = \begin{pmatrix} \boldsymbol{\Psi}_{1,1} & \cdots & \boldsymbol{\Psi}_{1,M} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Psi}_{M,1} & \cdots & \boldsymbol{\Psi}_{M,M} \end{pmatrix},$$

with $\boldsymbol{\Psi}_{m,m} = I_\nu$, $\boldsymbol{\Psi}_{m,m'} = \operatorname{cov}(Q(t_m), Q(t_{m'}))$.

- $\boldsymbol{T} = (T(t_1), \ldots, T(t_M))^\top$ follows a multivariate χ^2 distribution of Wishart type with the correlation matrix $\boldsymbol{\Psi}$ (Dickhaus and Royen, 2015).

Estimation of Ψ

- Consider SVD on $n^{-1}\boldsymbol{\Sigma}(t) = \boldsymbol{U}(t)\boldsymbol{V}(t)\boldsymbol{U}(t)^{\intercal}$.
  - Let $\boldsymbol{L}(t) = \boldsymbol{U}(t)\boldsymbol{V}^{1/2}(t)$ and $Q(t) = \boldsymbol{L}(t)^{\intercal}\boldsymbol{Z}(t)$
  - $T(t_m) = Q(t_m)^{\intercal}Q(t_m)$.
  - $\boldsymbol{\Psi}_{m,m'} = \boldsymbol{L}(t_m)^{\intercal}\,\mathrm{cov}(\boldsymbol{Z}(t_m), \boldsymbol{Z}(t_{m'}))\boldsymbol{L}(t_{m'}), t_m < t_{m'}$.
- We propose estimators of $\mathrm{cov}(\boldsymbol{Z}(t))$ and $\mathrm{cov}(\boldsymbol{Z}(t_m), \boldsymbol{Z}(t_{m'}))$ by asymptotic linearization.

Proposed Estimator of cov $(Z(t))$

- Since $d\bar{N}_{jkl}(s) = \sum_{i=1} W_{jkl,i}(s) dN_i(s)$, we have

$$Z_{jkl} = \sum_{i=1}^{n} \int_0^{\infty} \left\{ \frac{\bar{Y}_{111}(s) W_{jkl,i}(s)}{\bar{Y}_{jkl}(s) + \bar{Y}_{111}(s)} - \frac{\bar{Y}_{jkl}(s) W_{111,i}(s)}{\bar{Y}_{jkl}(s) + \bar{Y}_{111}(s)} \right\} dN_i(s).$$

- $\bar{Y}_{jkl}(s) = \sum_{i=1}^{n} W_{jkl,i}(s) Y_i(s)$ implies

$$\sum_{i=1}^{n} \left\{ \frac{\bar{Y}_{111}(s) W_{jkl,i}(s)}{\bar{Y}_{jkl}(s) + \bar{Y}_{111}(s)} - \frac{\bar{Y}_{jkl}(s) W_{111,i}(s)}{\bar{Y}_{jkl}(s) + \bar{Y}_{111}(s)} \right\} Y_i(s) = 0.$$

- Let $dM_i(s) = dN_i(s) - Y_i(s) d\Lambda_0(s)$ and $\pi_{jkl}(s) = \lim_{n\to\infty} \bar{Y}_{jkl}(s)/(\bar{Y}_{jkl}(s) + \bar{Y}_{111}(s))$, then

$$\frac{1}{\sqrt{n}} Z_{jkl} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_0^{\infty} \left\{ \frac{\bar{Y}_{111}(s) W_{jkl,i}(s)}{\bar{Y}_{jkl}(s) + \bar{Y}_{111}(s)} - \frac{\bar{Y}_{jkl}(s) W_{111,i}(s)}{\bar{Y}_{jkl}(s) + \bar{Y}_{111}(s)} \right\} dM_i(s)$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_0^{\infty} ((1-\pi_{jkl}(s)) W_{jkl,i}(s) - \pi_{jkl}(s) W_{111,i}(s)) dM_i(s) + o_p(1).$$

Proposed Estimator of cov $(\boldsymbol{Z}(t))$

- $Z_{jkl,i} = \int_0^\infty ((1-\pi_{jkl}(s))W_{jkl,i}(s) - \pi_{jkl}(s)W_{111,i}(s))dM_i(s)$ has mean 0 under H_0 based on Tsiatis and Davidian (2024).
- Let $\hat{\boldsymbol{Z}}_i = (\hat{Z}_{jkl,i})^\intercal \in \mathbb{R}^7$, where $\hat{Z}_{jkl,i} =$

$$\int_0^\infty \left\{ \frac{\bar{Y}_{111}(s)W_{jkl,i}(s)}{\bar{Y}_{jkl}(s)+\bar{Y}_{111}(s)} - \frac{\bar{Y}_{jkl}(s)W_{111,i}(s)}{\bar{Y}_{jkl}(s)+\bar{Y}_{111}(s)} \right\} \left\{ dN(s) - Y_i(s)\frac{dN(s)}{Y(s)} \right\}$$

  is the estimator of $Z_{jkl,i}$, with $dN(s)/Y(s)$ be the Nelson-Aalen estimator of $d\Lambda_0(s)$.
- Let $\hat{\boldsymbol{Z}}_i(t) = (\hat{Z}_{jkl,i}(t))^\intercal$ be $\hat{\boldsymbol{Z}}_i$ evaluated at time t, then $\text{cov}(\boldsymbol{Z}(t)/\sqrt{n})$ is estimated by

$$\frac{1}{n}\sum_{i=1}^{n} \hat{\boldsymbol{Z}}_i(t)\hat{\boldsymbol{Z}}_i(t)^\intercal.$$

Proposed Estimator of $\mathrm{cov}(\boldsymbol{Z}(t_m), \boldsymbol{Z}(t_{m'}))$

- $\mathrm{cov}(Z_{jkl,i}(t_m), Z_{jkl,i'}(t_{m'})) = 0$ for $i \neq i'$
- $\mathrm{cov}(Z_{jkl,i}(t_m), Z_{jkl,i}(t_{m'})) = E(Z_{jkl,i}(t_m) Z_{jkl,i}(t_{m'}))$ can be estimated by $Z_{jkl,i}(t_m) Z_{jkl,i}(t_{m'})$.
- An estimator of $\mathrm{cov}(\boldsymbol{Z}(t_m)/\sqrt{n_m}, \boldsymbol{Z}(t_{m'})/\sqrt{n_{m'}})$ when $t_m < t_{m'}$ is

$$\frac{1}{\sqrt{n_m n_{m'}}} \sum_{i=1}^{n_m} \hat{\boldsymbol{Z}}_i(t_m) \hat{\boldsymbol{Z}}_i(t_{m'})^{\mathsf{T}}.$$

Alternative Approaches for Boundary Calculation

- Full data is not available at the interim analysis time.
- Assume the independent increment property and $\text{cov}(Z(t_1))/n_1 \approx \text{cov}(Z(t_2))/n_2$ when we choose t_1 such that we expect to observe half of the events by time t_1.
- Specify error spending functions $\alpha^\star(t)$ and determine the boundaries sequentially (Lan and Demets, 1983).

Simulation Setting

- Simulated 1000 datasets; each with 500 patients arriving independently and uniformly over 5 years.
- Initial treatment assignment: $P\left(I_i\left(A_1\right)\right) = P\left(I_i\left(A_2\right)\right) = 0.5$.
- Second stage randomization indicator: $\eta_i \sim \text{Ber}(0.9)$.
- If $\eta_i = 0$, patient died in the 1st stage; time in stage 1 $T_{1i} \sim \text{Exp}\left(\boldsymbol{\theta}_j^N\right)$; survival time $T_i^S = T_{1i}$.
- If $\eta_i = 1$, we generated response indicator $R_i \sim \text{Ber}\left(\pi_R\right)$ and time in stage 1 $T_{1i} \sim \text{Exp}\left(\boldsymbol{\theta_j}\right)$.
  - If $R_i = 1$, patient i was a responder; maintenance treatment assignment: $P\left(I_i\left(B_1\right)\right) = P\left(I_i\left(B_2\right)\right) = 0.5$; time in stage 2 $T_{2i} \sim \text{Exp}\left(\boldsymbol{\theta}_{jk}^{R,E}\right)$; survival time $T_i^S = T_{1i} + T_{2i}$.
  - If $R_i = 0$, patient i was a non-responder; salvage treatment assignment: $P\left(I_i\left(C_1\right)\right) = P\left(I_i\left(C_2\right)\right) = 0.5$; time in stage 2 $T_{2i} \sim \text{Exp}\left(\boldsymbol{\theta}_{jl}^{NR,E}\right)$; survival time $T_i^S = T_{1i} + T_{2i}$.
- Censoring time: $C_i \sim \text{Unif}(0, 5)$.

Simulation

- Null Scenario: $\theta_j = \theta_j^N = (5,5), \theta_{jk}^{R,E} = (5,5,5,5), \theta_{jl}^{NR,E} = (5,5,5,5)$.

M	Boundary	Rejection Rate	LR	BLR	TD	BTD
1	NA	Type I Error	0.038	0.035	0.061	0.069
2	Pocock	Rej at interim	0.022	0.029	0.040	0.049
		Rej at final	0.015	0.016	0.027	0.027
		Type I Error	0.037	0.045	0.066	0.075
	OBF	Rej at interim	0.005	0.007	0.017	0.022
		Rej at final	0.038	0.036	0.050	0.066
		Type I Error	0.043	0.043	0.066	0.087
	LD-Pocock	Rej at interim	0.030	0.038	0.046	0.062
		Rej at final	0.014	0.014	0.024	0.027
		Type I Error	0.044	0.051	0.069	0.087
	LD-OBF	Rej at interim	0.003	0.006	0.014	0.015
		Rej at final	0.037	0.032	0.051	0.067
		Type I Error	0.040	0.038	0.064	0.081

Table 1: False Rejection Rates under Null Scenario over 1000 simulations.

Simulation

M	Boundary	Rejection Rate	Alt1		Alt2	
			LR	TD	LR	TD
1	NA	Power	0.78	0.82	0.92	0.95
2	Pocock	Rej at interim	0.32	0.40	0.45	0.47
		Rej at final	0.59	0.64	0.78	0.81
		Power	0.72	0.78	0.88	0.92
		E(n)	434	415	406	380
	OBF	Rej at interim	0.17	0.27	0.27	0.45
		Rej at final	0.74	0.76	0.89	0.91
		Power	0.79	0.83	0.92	0.95
		E(n)	465	443	443	406
	LD-Pocock	Rej at interim	0.34	0.43	0.49	0.62
		Rej at final	0.57	0.63	0.77	0.89
		Power	0.72	0.79	0.88	0.92
		E(n)	428	410	399	371
	LD-OBF	Rej at interim	0.16	0.24	0.26	0.40
		Rej at final	0.73	0.77	0.88	0.91
		Power	0.78	0.82	0.91	0.95
		E(n)	466	450	445	416

Table 2: False Rejection Rates under Alt Scenario over 1000 simulations.

Discussion and Future Work

- We proposed an interim analysis procedure for SMART for survival outcomes, and also provided the statistical properties of the proposed test statistics.
- IM-SMART can potentially end the trial early without inflating type I error and maintaining statistical power.
- May extend the proposed method to various design settings.

Thank You!

References I

T. Dickhaus and T. Royen. A survey on multivariate chi-square distributions and their applications in testing multiple hypotheses. *Statistics*, 49(2):427–454, 2015.

M. Gu and T. L. Lai. Determination of power and sample size in the design of clinical trials with failure-time endpoints and interim analyses. *Controlled Clinical Trials*, 20:423–438, 1999.

K. Kim and A. Tsiatis. Study duration for clinical trials with survival response and early stopping rule. *Biometrics*, 46(1):81–92, 1990.

K. Lan and D. L. Demets. Discrete sequential boundaries for clinical trials. *Biometrics*, 70: 659–663, 1983.

C. Manschot, E. Laber, and M. Davidian. Interim monitoring of sequential multiple assignment randomized trials using partial information. *Biometrics*, 79:2881–2894, 2023.

P. C. O'Brien and T. R. Fleming. A multiple testing procedure for clinical trials. *Biometrics*, 35 (3):549–556, 1979.

S. J. Pocock. Group sequential methods in the design and analysis of clinical trials. *Biometrika*, 64:191–199, 1977.

Y. Shen and J. Cai. Sample size reestimation for clinical trials with censored survival data. *Journal of the American Statistical Association*, 98(462):418–426, 2003.

A. Tsiatis. Repeated significance testing for a general class of statistics used in censored survival analysis. *Journal of the American Statistical Association*, 77(380):855–861, 1982.

References II

A. A. Tsiatis and M. Davidian. A generalized logrank-type test for comparison of treatment regimes in sequential multiple assignment randomized trials. *Biometrics. arXiv preprint arXiv:2403.16813*, 2024.

L. Wu, J. Wang, and A.S. Wahed. Interim monitoring in sequential multiple assignment randomized trials. *Biometrics*, 79(1):368–380, 2023.

Parameters under Alternative Scenarios

- Alternative Scenario 1:
  $\boldsymbol{\theta}_j = \boldsymbol{\theta}_j^N = (5,5), \boldsymbol{\theta}_{jk}^{R,E} = (2,4,3,4), \boldsymbol{\theta}_{jl}^{NR,E} = (3.2,2,2.9,2).$
- Alternative Scenario 2: $\boldsymbol{\theta}_j = \boldsymbol{\theta}_j^N = (5,5)$,
  $\boldsymbol{\theta}_{jk}^{R,E} = (2.8,4.6,2.3,4.9), \boldsymbol{\theta}_{jl}^{NR,E} = (5.8,4.3,5.2,6.5).$

Details of Lan-Demets Boundaries

- $\alpha^\star(t)$ is a strictly increasing function; boundaries b_1 and b_2 satisfy $\text{pr}(T(t_1) > b_1) = \alpha^\star(t_1)$ and $\text{pr}(T(t_1) \leq b_1, T(t_2) > b_2) = \alpha - \alpha^\star(t_1)$.
- Choose $\alpha_1^\star(t) = 2(1 - \Phi(z_{\alpha/2}/(t/L)^{1/2})$ and $\alpha_2^\star(t) = \alpha \log(1 + (e-1)t/L)$ to generate boundaries that are similar to the Pocock and OBF type boundaries with $L = 2t_1$, where Φ is the cumulative distribution function of a standard normal random variable.