Sequential Monitoring of Covariate Adaptive Randomized Clinical Trials with Sample Size Re-estimation

Jun Yu[1], Dejian Lai[2]

[1] AbbVie Inc., North Chicago, Illinois, United States; [2]University of Texas Health Science Center at Houston, Houston, Texas, United States

Presented at the ASA Biopharmaceutical Section Regulatory-Industry Statistics Workshop, September 22-25, 2020, Washington, DC

ABSTRACT

Once scientific questions are determined, other design features of clinical trials including increasing the power while controlling the type I error rate, planning interim analysis, and achieving treatment balance among subgroups will be either required or preferred. We propose to sequentially monitor the covariate adaptive randomization (CAR) procedures with sample size re-estimation (SSR) to satisfy a variety of design objectives of clinical trials. However, each of the three adaptive designs (sequential monitoring, CAR, and SSR) poses a challenge to the control of the type I error rate. In this research, we investigated how to utilize the advantages of the three adaptive methods and control the type I error rate. We proved that the asymptotic joint distribution of the sequential statistics follows the asymptotic canonical joint distribution defined in Jennison and Turnbull (2000). Besides, numerical studies demonstrated that our methods could control the type I error rate, increase the power, and lead to much-improved treatment balance across subgroups.

INTRODUCTION

- Covariate adaptive randomization (CAR), sequential monitoring, and sample size re-estimation (SSR) are often preferred in practice.
- However, there is not a comprehensive theoretical and numerical study about sequential monitoring of CAR procedures with SSR in the literature.
- Consequence: The clinical trialists usually implement sequential monitoring of CAR with SSR based on the theories on complete randomization and assume the discrepancy is negligible.

FRAMEWORK and ASYMPTOTIC RESULTS

Consider a two-arm randomized clinical trial with CAR,

- Originally planned sample size: n
- Treatment assignments: T_i $(i = 1, \ldots, n)$
- Covariates: $\boldsymbol{X}_i = (\boldsymbol{X}_{i1}, \ldots, \boldsymbol{X}_{ip})$ and $\boldsymbol{V}_i = (\boldsymbol{V}_{i1}, \ldots, \boldsymbol{V}_{iq})$
- Responses: Y_i $(i = 1, \ldots, n)$

$$Y_i = \mu_1 T_i + \mu_2(1 - T_i) + X_{i1}\beta_1 + \ldots + X_{ip}\beta_p + V_{i1}\gamma_1 + \ldots + V_{iq}\gamma_q + \epsilon_i, \quad (1)$$

where

- μ_1 and μ_2 are treatment effects for treatment 1 and 2,
- ϵ_i are independent errors with mean 0 and variance σ^2.
- Assume that the covariates $(X_1, \ldots, X_p)$ and $(V_1, \ldots, V_q)$ are used to implement CAR

We perform the following hypothesis testing to compare two treatments in clinical trials:

$$H_0 : \mu_1 = \mu_2 \text{ versus } \mu_1 \neq \mu_2. \quad (2)$$

Sequential monitoring with sample size re-estimation:

- K interim analyses at information time points $t_1, \ldots, t_L, \ldots, t_K$.
- We implement SSR at the end of the Lth interim analysis $(L < K)$ based on the observed data using the method in Cui et al. (1999).
- We first calculate the conditional power, CP_L, based on observed data for originally planned sample size n. If CP_L is not less than the desirable level of cp, then no SSR will be implemented. Otherwise, search n^* that satisfies $CP_L = cp$.
- Next, we increase the original sample size at stages $k \geq L + 1$ by a multiplier of $b = \min(b^*, b_{\max})$, where $b_{\max}$ is a prespecified maximum sample size factor, and $b^* = (n^* - N)/(n - N)$.

Write $\boldsymbol{\mu} = (\mu_1, \mu_2)^T$, $\boldsymbol{\eta} = (\mu_1, \mu_2, \beta_1, \ldots, \beta_p)^T$, $\boldsymbol{T}(n) = (T_1, \ldots, T_n)^T$, $\boldsymbol{Y}(n) = (Y_1, \ldots, Y_n)^T$, $\boldsymbol{\epsilon}(n) = (\epsilon_1, \ldots, \epsilon_n)^T$, $\boldsymbol{L} = (1, -1, 0, ..., 0)$, and

$$\boldsymbol{X}(n) = \begin{bmatrix} T_1 & 1 - T_1 & X_{11} & \ldots & X_{1p} \\ T_2 & 1 - T_2 & X_{21} & \ldots & X_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ T_n & 1 - T_n & X_{n1} & \ldots & X_{np} \end{bmatrix}.$$

When implementing CAR regarding continuous covariates, we first discretize these continuous covariates and apply CAR designs regarding the discretized covariates. Specifically, let

$$\tilde{\boldsymbol{X}}_k = \begin{cases} \boldsymbol{X}_k & \text{if } k \notin C \\ d_k(\boldsymbol{X}_k) & \text{if } k \in C \end{cases}$$

FRAMEWORK and ASYMPTOTIC RESULTS (CONTINUED)

and

$$\tilde{\boldsymbol{V}}_j = \begin{cases} \boldsymbol{V}_j & \text{if } j \notin C^* \\ d_j^*(\boldsymbol{V}_j) & \text{if } j \in C^* \end{cases},$$

where $C = \{l | \boldsymbol{X}_l \text{ is continuous}, l = 1, \ldots, p\}$, $C^* = \{l | \boldsymbol{V}_l \text{ is continuous}, l = 1, \ldots, q\}$, and $d(\cdot)$ and $d^*(\cdot)$ are certain discrete functions.

- We propose the following sequential statistics

$$U_t = \begin{cases} Z_t^{adj}, & \text{if } t \leq t_L \\ {w_t}^{1/2} \times Z_{t_L}^{adj} + (1 - w_t)^{1/2} \times \\ \quad \{[B(b(t - t_L) + t_L) - B(t_L)]/[b(t - t_L)]^{1/2}\}, & \text{if } t > t_L \end{cases} \quad (3)$$

where $w_t = t_L/t$, $B(t) = \sqrt{t} Z_t^{adj}$,

$$Z_t^{adj} = \frac{\boldsymbol{L}\hat{\boldsymbol{\eta}}(t)}{\hat{\epsilon}(t)\sqrt{\hat{\sigma}(t)^2 \boldsymbol{L}(\boldsymbol{X}(\lfloor nt \rfloor)^T \boldsymbol{X}(\lfloor nt \rfloor))^{-1} \boldsymbol{L}^T}}, \quad (4)$$

and $\hat{\epsilon}(t)^2$ is any consistent estimator of

$$\frac{\sum_{j \in C^*} \gamma_j^2 \sigma_{\delta j}^2 + \sigma^2}{\sigma^2 + \sum_{j=1}^{q} Var(V_j \gamma_j^T)}, \quad (5)$$

$\sigma_{\delta j}^2 = E\left[Var\left(\delta_j | d_j^*(V_j)\right)\right]$, and $\delta_j = V_j - E(V_j | d_j^*(V_j))$.

Let $W_i = (x_{i1}^{c_1}, \ldots, x_{ip}^{c_p}, v_{i1}^{c_1^*}, \ldots, v_{iq}^{c_q^*})$ represent the ith subject's covariate profile if $\tilde{\boldsymbol{X}}_{ik}$ is at level $x_{ik}^{c_k}$ and $\tilde{\boldsymbol{V}}_{ij}$ is at level $v_{ij}^{c_j^*}$.

- DIF_n: the overall difference in patient numbers between two treatments after n patients.
- $DIF_n(k; c_k)$: the marginal difference with respect to the level $x_k^{c_k}$ of covariate X_k.
- $DIF_n(j; c_j^*)$: the marginal difference with respect to the level $v_j^{c_j^*}$ of covariate $\tilde{\boldsymbol{V}}_j$.
- $DIF_n(c_1, \ldots, c_p, , c_1^*, \ldots, c_q^*)$: the difference in patient numbers in the stratum containing the subjects with covariates $(x_1^{c_1}, \ldots, x_p^{c_p}, v_1^{c_1^*}, \ldots, v_q^{c_q^*})$.

Theorem 1 *Let $B_t^U = \sqrt{t} U_t$. Assume the CAR design satisfies $DIF_n = O_p(1)$, $DIF_n^X(k; c_k) = O_p(1), k = 1, \ldots, p$, and $DIF_n^V(j; c_j^*) = O_p(1), j = 1, \ldots, q$. Then under H_0, B_t^U is asymptotically a standard Brownian motion in distribution, and the sequence of test statistics $\{(U_{t_1}, \ldots, U_{t_K}), 0 \leq t_1 \leq t_2 \leq \ldots \leq t_K \leq 1\}$ satisfies*
(i) $\{U_{t_1}, ..., U_{t_K}\}$ follows multivariate normal distribution;
(ii) $EU_{t_i} = 0$;
(iii) $Cov(U_{t_i}, U_{t_j}) = \sqrt{\lfloor nt_i \rfloor / \lfloor nt_j \rfloor}$, $0 \leq t_i \leq t_j \leq 1$.

Two special cases:

- (1) If all the randomization covariates are used in the data analysis, and

$$Y_i = \mu_1 T_i + \mu_2(1 - T_i) + X_{i1}\beta_1 + \ldots + X_{ip}\beta_p + \epsilon_i. \quad (6)$$

then we do not need to adjust the regular statistics obtained from fitting the linear regression and use

$$Z_t = \frac{L\hat{\boldsymbol{\eta}}(t)}{\sqrt{\hat{\sigma}(t)^2 L(\boldsymbol{X}(\lfloor nt \rfloor)^T \boldsymbol{X}(\lfloor nt \rfloor))^{-1} L^T}}, \quad (7)$$

to calculate U_t.

- (2) If we do not want to include any randomization covariates in the data analysis (t-test), we just need to divide the t-test statistics by $\hat{\epsilon}(t)$ to obtain Z_t^{adj} for our procedure.

NUMERICAL STUDIES

$$Y_i = \mu_1 T_i + \mu_2(1 - T_i) + X_{i1}\beta_1 + X_{i2}\beta_2 + \epsilon_i, i = 1, \ldots, 500, \quad (8)$$

Equivalently, it can be written as

$$Y_i = \beta_0 + \beta_T T_i + X_{i1}\beta_1 + X_{i2}\beta_2 + \epsilon_i, i = 1, \ldots, 500. \quad (9)$$

- Originally planned sample size: $n = 500$
- $(\beta_1, \beta_2) = (1, 1)$
- ϵ_i are independent errors from normal distribution $N(0, 1)$
- Case 1: $X_1 \in Bernoulli(p_1)$ and $X_2 \in Bernoulli(p_2)$.
- Case 2: $X_1 \in N(0, 1)$ and $X_2 \in N(0, 1)$.
- Time points: $t_1 = 0.2$ $(n_1 = 100)$, $t_2 = 0.5$ $(n_2 = 250)$, $t_3 = 1$ $(n_3 = 500)$.
- O'Brien-Fleming-like boundaries $(4.877, 2.963, 1.969)$.

NUMERICAL STUDIES (CONTINUED)

- We implement SSR if the trial is determined to continue after the second interim analysis.
- The cap of the sample size at stage 3 is 500 and $b_{\max} = 2$.
- All the results are based on $10,000$ replications.

Table 1: Performance of different designs under H_0 when all the randomization covariates are included in the data analysis.

(β_0, p_1, p_2)	Design	α	$\hat{\beta}_T$	DIF_n	DIF_{11}	$DIF_{1.}$
Both covariates are discrete						
$(0.5, 0.5, 0.5)$	SPB	0.049	0.000(0.080)	1.32(1.27)	0.67(0.62)	0.99(0.84)
$(0.5, 0.5, 0.5)$	PS	0.051	0.000(0.080)	1.70(1.70)	5.44(4.16)	1.54(1.42)
$(0.5, 0.5, 0.5)$	CR	0.051	0.001(0.080)	20.7(15.7)	10.4(7.91)	14.5(11.1)
$(2, 0.4, 0.6)$	SPB	0.050	0.000(0.079)	1.33(1.27)	0.66(0.62)	0.99(0.83)
$(2, 0.4, 0.6)$	PS	0.051	0.000(0.081)	1.68(1.66)	5.22(4.00)	1.52(1.41)
$(2, 0.4, 0.6)$	CR	0.050	0.000(0.079)	20.8(16.2)	10.3(7.84)	13.3(10.1)
Both covariates are continuous						
(β_0, q_1, q_2)	Design	α	$\hat{\beta}_T$	DIF_n	DIF_{11}	$DIF_{1.}$
$(0.5, 0.5, 0.5)$	SPB	0.055	0.000(0.081)	1.32(1.28)	0.68(0.62)	1.00(0.84)
$(0.5, 0.5, 0.5)$	PS	0.049	-0.001(0.080)	1.69(1.68)	5.43(4.14)	1.53(1.38)
$(0.5, 0.5, 0.5)$	CR	0.055	0.000(0.082)	20.8(16.0)	10.5(8.04)	14.8(11.4)
$(2, 0.4, 0.6)$	SPB	0.054	0.000(0.080)	1.33(1.27)	0.66(0.62)	0.99(0.83)
$(2, 0.4, 0.6)$	PS	0.052	0.001(0.081)	1.71(1.69)	5.17(3.95)	1.52(1.39)
$(2, 0.4, 0.6)$	CR	0.057	0.000(0.082)	21.2(16.0)	10.0(7.79)	13.2(10.0)

Note: α: type I error rate; $\hat{\beta}_T$: estimator of β_T; DIF_n: overall difference in patient numbers between the two treatments; DIF_{11}: the difference of patient numbers between the two treatments in the stratum formed by $X_1 = 1$ and $X_2 = 1$; $DIF_{1.}$: marginal imbalance for $X_1 = 1$.

Table 2: Performance of CAR under H_0 when unadjusted t-test is used.

(β_0, p_1, p_2)	X_1, X_2	Design	α	(β_0, q_1, q_2)	X_1, X_2	Design	α
$(0.5, 0.5, 0.5)$	discrete	SPB	0.013	$(0.5, 0.5, 0.5)$	continuous	SPB	0.006
$(0.5, 0.5, 0.5)$	discrete	PS	0.015	$(0.5, 0.5, 0.5)$	continuous	PS	0.007
$(0.5, 0.4, 0.6)$	discrete	SPB	0.013	$(0.5, 0.4, 0.6)$	continuous	SPB	0.008
$(0.5, 0.4, 0.6)$	discrete	PS	0.014	$(0.5, 0.4, 0.6)$	continuous	PS	0.010
$(2, 0.5, 0.5)$	discrete	SPB	0.010	$(2, 0.5, 0.5)$	continuous	SPB	0.007
$(2, 0.5, 0.5)$	discrete	PS	0.016	$(2, 0.5, 0.5)$	continuous	PS	0.006
$(2, 0.4, 0.6)$	discrete	SPB	0.011	$(2, 0.4, 0.6)$	continuous	SPB	0.006
$(2, 0.4, 0.6)$	discrete	PS	0.012	$(2, 0.4, 0.6)$	continuous	PS	0.007

Table 3: Performance of CAR under H_0 when adjusted t-test is used.

(β_0, p_1, p_2)	Design	α	DIF_n	DIF_{11}	$DIF_{1.}$
Both covariates are discrete					
$(0.5, 0.5, 0.5)$	SPB	0.052	1.32(1.27)	0.67(0.62)	0.99(0.84)
$(0.5, 0.5, 0.5)$	PS	0.051	1.68(1.69)	5.43(4.16)	1.51(1.37)
$(0.5, 0.5, 0.5)$	CR	0.049	20.8(15.8)	10.3(7.90)	14.6(11.1)
$(2, 0.4, 0.6)$	SPB	0.049	1.33(1.27)	0.66(0.62)	0.99(0.83)
$(2, 0.4, 0.6)$	PS	0.055	1.68(1.68)	5.21(3.98)	1.52(1.38)
$(2, 0.4, 0.6)$	CR	0.051	20.9(16.1)	10.3(7.83)	13.3(10.1)
Both covariates are continuous					
(β_0, q_1, q_2)	Design	α	DIF_n	DIF_{11}	$DIF_{1.}$
$(0.5, 0.5, 0.5)$	SPB	0.050	1.33(1.27)	0.68(0.62)	0.98(0.84)
$(0.5, 0.5, 0.5)$	PS	0.055	1.70(1.67)	5.45(4.18)	1.54(1.39)
$(0.5, 0.5, 0.5)$	CR	0.051	20.8(16.0)	10.5(8.01)	14.8(11.4)
$(2, 0.4, 0.6)$	SPB	0.054	1.32(1.28)	0.66(0.63)	0.99(0.83)
$(2, 0.4, 0.6)$	PS	0.056	1.70(1.68)	5.20(3.96)	1.53(1.42)
$(2, 0.4, 0.6)$	CR	0.048	21.2(16.1)	10.0(7.77)	13.2(10.0)

Note: α: type I error rate; DIF_n: overall difference in patient numbers between the two treatments; DIF_{11}: the difference of patient numbers between the two treatments in the stratum formed by $X_1 = 1$ and $X_2 = 1$; $DIF_{1.}$: marginal imbalance for $X_1 = 1$.

Table 4: Performance of CAR under H_1 when adjusted t-test is used.

(β_T, p_1, p_2)	Design	Power	DIF_n	DIF_{11}	$DIF_{1.}$
Both covariates are discrete					
$(0.3, 0.5, 0.5)$	SPB	0.893	1.33(1.26)	0.65(0.62)	0.99(0.84)
$(0.3, 0.5, 0.5)$	PS	0.892	1.69(1.67)	4.46(3.54)	1.53(1.40)
$(0.3, 0.5, 0.5)$	CR	0.735	18.2(14.3)	9.21(7.16)	12.9(10.1)
$(0.35, 0.4, 0.6)$	SPB	0.960	1.30(1.25)	0.67(0.63)	1.00(0.83)
$(0.35, 0.4, 0.6)$	PS	0.965	1.71(1.67)	4.01(3.20)	1.54(1.39)
$(0.35, 0.4, 0.6)$	CR	0.870	17.4(13.7)	8.48(6.71)	11.0(8.66)
Both covariates are continuous					
(β_T, q_1, q_2)	Design	Power	DIF_n	DIF_{11}	$DIF_{1.}$
$(0.4, 0.5, 0.5)$	SPB	0.903	1.32(1.27)	0.67(0.62)	0.99(0.84)
$(0.4, 0.5, 0.5)$	PS	0.900	1.71(1.69)	4.46(3.55)	1.55(1.41)
$(0.4, 0.5, 0.5)$	CR	0.691	18.6(14.6)	9.36(7.22)	13.3(10.3)
$(0.45, 0.4, 0.6)$	SPB	0.953	1.30(1.26)	0.67(0.63)	1.00(0.84)
$(0.45, 0.4, 0.6)$	PS	0.948	1.70(1.67)	4.11(3.29)	1.54(1.40)
$(0.45, 0.4, 0.6)$	CR	0.792	18.2(14.2)	8.82(6.80)	11.3(8.86)

Note: α: type I error rate; DIF_n: overall difference in patient numbers between the two treatments; DIF_{11}: the difference of patient numbers between the two treatments in the stratum formed by $X_1 = 1$ and $X_2 = 1$; $DIF_{1.}$: marginal imbalance for $X_1 = 1$.

CONCLUSIONS

- When unadjusted t-test is used, the type I error rates are all conservative for CAR when both covariates are either discrete or continuous.
- When all the randomization covariates are included in the data analysis or adjusted t-test is used, our methods can control the type I error rate well and accurately estimate the unknown parameters.
- CAR can lead to significantly better overall, marginal or stratum-level balance than complete randomization.
- Our proposed methods can increase the power.