



# Article Estimating Conditional Power for Sequential Monitoring of Covariate Adaptive Randomized Designs: The Fractional Brownian Motion Approach

Yiping Yang, Hongjian Zhu and Dejian Lai \*

Department of Biostatistics and Data Science, School of Public Health, The University of Texas Health Science Center at Houston, 1200 Herman Pressler Street, Room 1008, Houston, TX 77030, USA; dieling414@gmail.com (Y.Y.); hongjian.zhu@uth.tmc.edu (H.Z.)

\* Correspondence: dejian.lai@uth.tmc.edu; Tel.: +1-(713)-500-9270

**Abstract:** Conditional power based on classical Brownian motion (BM) has been widely used in sequential monitoring of clinical trials, including those with the covariate adaptive randomization design (CAR). Due to some uncontrollable factors, the sequential test statistics under CAR procedures may not satisfy the independent increment property of BM. We confirm the invalidation of BM when the error terms in the linear model with CAR design are not independent and identically distributed. To incorporate the possible correlation structure of the increment of the test statistic, we utilize the fractional Brownian motion (FBM). We conducted a comparative study of the conditional power under BM and FBM. It was found that the conditional power under FBM assumption was mostly higher than that under BM assumption when the Hurst exponent was greater than 0.5.

**Keywords:** Hurst exponent; covariate-adaptive randomization trial; sequential monitoring; Brownian motion; fractional Brownian motion; conditional power; drift parameters

## 1. Introduction

Clinical trials, aiming to evaluate the safety and efficacy of drugs and medical devices in target populations, play an important role in the development of public health and medicine [1]. Adaptive randomized clinical design improves the trial, based on the accumulated data and changing environment, making clinical trials more efficient, flexible, and ethically reasonable [2].

The treatment effects estimated from unadjusted models may not be appropriate when the covariates are not balanced. Meanwhile, some covariates, such as elevated values of biomarkers that were found to affiliate with disease status in the translational research, may be critical in determining the treatment effects in clinical trials [3]. For example, biomarker HR23B is closely related to and used to indicate the effectiveness of histone deacetylase inhibitors-based therapy for tumors [4]. To address the problem of covariate imbalance, a useful tool is the Covariate Adaptive Randomization (CAR) procedure in which participants are assigned to different treatment groups based on previous participants' assignment, previous participants' covariates, and current participants' covariates, such that the asymmetries across the subgroups are minimized [5]. Although the complete randomization is good at eliminating the selection bias, the CAR design is a more reasonable comprise between reducing the selection bias and balancing covariates assignments [6]. The rigorous theory of covariate adaptive randomized clinical trials has been developed more recently [7]. Thereafter, progress has been made in research on statistical inference with CAR designs [8,9].

Traditionally, classical Brownian motion (BM) is a fundamental theory for monitoring outcome effects in clinical trials, including those with CAR designs [10–14]. It has been proved that the sequential test statistics of covariate adaptive clinical trials follow Brownian motion asymptotically under some regularized conditions [15].



Citation: Yang, Y.; Zhu, H.; Lai, D. Estimating Conditional Power for Sequential Monitoring of Covariate Adaptive Randomized Designs: The Fractional Brownian Motion Approach. *Fractal Fract.* **2021**, *5*, 114. https://doi.org/10.3390/ fractalfract5030114

Academic Editors: Manuel Ortigueira, Duarte Valério and Palle Jorgensen

Received: 12 August 2021 Accepted: 5 September 2021 Published: 8 September 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). A condition of performing the hypothesis testing of covariate adaptive randomized clinical trials is that the underlying error terms are independent and identically distributed (iid). In addition, independent increments are a property of classical Brownian motion [16]. However, the independent increment property of the test statistics may not be completely met in some situations. For example, some patients may enter the trial during the same season; some patients may be treated by the same hospital or the same physician. Therefore, the error terms from the model may be correlated and follow special covariance patterns. Given the situation that most of the previous theoretical research into sequential monitoring of CAR designs was based on the Brownian motion assumption, it is necessary to explore the stochastic properties of the sequential monitoring process when error structures are not independent and identically distributed. We propose fractional Brownian motion (FBM) as a more valid tool to investigate the outcomes of clinical trials with correlated error structures.

FBM, annotated as " $B_H(t)$ ", is a Gaussian process with  $E(B_H(t)) = 0$ , and  $Cov(B_H(s), B_H(t)) = \frac{1}{2}(t^{2H} + s^{2H} - |t - s|_{2H})$ , where the Hurst coefficient (H), in the range of 0 < H < 1, is a parameter of the FBM, describing the long-term dependence of the process [17–19]. FBM is a Markov process when H = 1/2 [20,21]. The maximum likelihood estimation (MLE) method was proposed for estimating the Hurst coefficient underlying FBM in clinical trials [22]. The log likelihood function of the observed value of  $B_H = (B_H(t_1), \dots, B_H(t_n))^t$  was  $L_n(B_H, H) = -nlog(2\pi)/2 - log|\Sigma(H)|/2 - B_H^t \Sigma^{-1}(H)B_H/2$ , where  $\Sigma(H)$  is the variance covariance matrix of  $B_H$  [22].

In this paper, comprehensive simulations of the sequential monitoring of CAR procedures were conducted to investigate the breakdown of BM when the independent increment assumption was not met. We further calculated the conditional power (CP) under the null hypothesis, with the BM assumption and with the FBM assumption, respectively.

Section 2 of this paper includes the test statistics and theoretical properties under covariate adaptive randomized clinical trials with correlated error structures. In Section 3, results from numerical simulations are provided to estimate the Hurst exponents for sequential monitoring of CAR procedures when error structures are not iid. Conditional powers are calculated and compared between BM and FBM assumptions in Section 4. Conclusions and discussions are found in Section 5.

#### 2. Covariate-Adaptive Randomized Clinical Trials with Correlated Error Structures

Assume there is a two-arm randomized sequential clinical trial with covariate adaptive designs.  $\mu_1$  and  $\mu_2$  are parameters for group 1 and group 2, measuring the main effects of the treatment in each group, respectively.  $I_i$  are the indicator variables for *i*th patient assigned to the different treatment groups, i = 1, 2, ...N. When  $I_i = 1$ , the patient is assigned to the treatment group 1; when  $I_i = 0$ , the patient is assigned to the treatment group 2. Let  $Y_i$  be the response outcomes of the model.  $X_{i, 1} ... X_{i, p}$  are covariates.  $\beta_1 ... \beta_p$  are unknown parameters for the covariates. p is the number of covariates besides  $\mu_1$  and  $\mu_2$ . The expectation of X is equal to 0. All covariates are independent from each other.  $\varepsilon_i$ /s denote the error terms that are correlated. N is the total sample size.

$$Y_{i} = \mu_{1}I_{i} + \mu_{2}(1 - I_{i}) + \beta_{1}X_{i, 1} + \dots + \beta_{p}X_{i, p} + \varepsilon_{i} \ i = 1, \dots, N$$
(1)  
$$E(X_{k}) = 0, \text{ for } k = 1, 2, \dots, p$$

 $X_{i, k}$ 's are covariates independent with the error terms  $\varepsilon_i$ . Indicator  $I_i$  are also independent with the error terms  $\varepsilon_i$ .

Let 
$$Y = (Y_1, Y_2, \ldots, Y_N)^T$$
,  $\beta = (\mu_1, \mu_2, \beta_1, \ldots, \beta_p)^T$ ,  $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_N)^T$ 

$$X = \begin{bmatrix} I_1 & 1 - I_1 & X_{1,1} & \cdots & X_{1,p} \\ I_2 & 1 - I_2 & X_{2,1} & \cdots & X_{2,p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ I_N & 1 - I_N & X_{N,1} & \cdots & X_{N,p} \end{bmatrix}$$

The expression (1) can be synthesized to the matrix form:  $Y = X\beta + \varepsilon$ . The hypothesis for testing the treatment effect is:

$$H_0: \mu_1 - \mu_2 = 0$$
 vs.  $H_A: \mu_1 - \mu_2 \neq 0$ 

When error terms are independent and normally distributed, the statistics for testing the main outcome effect at the time point t can be denoted as [7]:

$$Z_t^* = \frac{L\hat{\beta}^*(t)}{\sqrt{\hat{\sigma}^*(t)^2 L(X(\lfloor Nt \rfloor)^T X(\lfloor Nt \rfloor))^{-1} L^T}}$$
(2)

$$B_{H}^{*}(t) = \frac{\sqrt{tL}\hat{\beta}^{*}(t)}{\sqrt{\hat{\sigma}^{*}(t)^{2}L(X(\lfloor Nt \rfloor)^{T}X(\lfloor Nt \rfloor))^{-1}L^{T}}}$$
(3)

$$\hat{\beta}^*(t) = \left(X(\lfloor Nt \rfloor)^T X(\lfloor Nt \rfloor)\right)^{-1} X\left(\lfloor Nt \rfloor\right)^T Y(\lfloor Nt \rfloor)$$
(4)

$$\hat{\sigma}^{2*}(t) = \left(Y(\lfloor Nt \rfloor) - X(\lfloor Nt \rfloor)\hat{\beta}^*(t)\right)^T \left(Y(\lfloor Nt \rfloor) - X(\lfloor Nt \rfloor)\hat{\beta}(t)^*\right) / \left(\lfloor Nt \rfloor - p - 2\right)$$
(5)

Under null hypothesis with equal treatment effect, we have:

$$L = (1, -1, 0, \dots, 0)t = n/N, n < N, t \in [0, 1]$$

 $|\cdot|$  is denoted as the floor function.

When using the statistic (2), the sequential monitoring process with multiple interim analyses is considered. For example, the investigator may evaluate the test statistics after every 100 patients finish the study in a 4000 patients' clinical trial.  $Z_1$  is the statistic test for patients 1–100;  $Z_2$  is for patients 1–200; ...;  $Z_{40}$  is for patients 1–4000. Each test statistic Z value is different from others, with different recruited patient numbers at that time point.

It was indicated that the test statistic *Z* score at each time point in (2) followed normal distribution under the null hypothesis [7]. Interim normalized *Z* value can be transformed into *B*-value  $B_{(H)}$  as  $B_H(t) = Z_t t^{1/2}$ . The sequential statistics test  $B_H$  in the model (3) converges to asymptotically Brownian motion [12]. Both conclusions above were based on the assumption of error terms iid in the model. In some cases, the independent increment property of the test statistic may be invalid when error terms  $\varepsilon$ 's of linear model (1) are correlated. Nonetheless, people may ignore the covariance in the error terms and still use the original classical hypothesis test, with error terms independent and identically distributed. This scenario is called misspecification in our study. \* used in the above Equations (2)–(5) were the mis-specified test equations and parameters, distinguished from the classical original hypothesis test formula when  $\varepsilon$ 's are iid.

If the sequential test statistics  $B_H^*(t)$  converge to a standard Brownian motion,  $\operatorname{Cov}(Z_{t_i}^*, Z_{t_j}^*) = \sqrt{t_i/t_j} \ 0 \le t_i \le t_j \le 1$  [23]. Then  $\operatorname{Cov}(B_H^*(t_i), B_H^*(t_j)) \xrightarrow{p} t_i$  since  $B_H^*(t) = Z_t^* \sqrt{t}$ . However, based on the theoretical derivation in Yang [15], the covariance of  $B_H^*(t_i)$  and  $B_H^*(t_j)$  could not converge to the minimum of  $t_i$  and  $t_j$  under the null hypothesis in misspecification scenarios. The conclusion demonstrated that  $B_H^*(t)$  in sequential test statistics (3) was not asymptotically standard Brownian motion under the misspecification scenarios.

#### 3. Simulations for Misspecification Scenarios

Since the sequential test statistics cannot converge to asymptotically Brownian motion when error terms are correlated, we propose a larger class of fractional Brownian motion for the stochastic structure of the test statistic. Maximum likelihood method was used to estimate the Hurst exponents of the FBM for the sequential monitoring processes under the misspecification assumption. If the mean estimation of H values deviates significantly from 0.5, the sequential monitoring processes would be confirmed as not converging to BM. Error terms in the model (1) were assumed to follow specific correlated patterns. Increments of fractional Brownian motion, defined as  $W_i = B_H(t_i) - B_H(t_{i-1})$ , were used in the error terms  $\varepsilon$  of our simulations [24]. fbm() function in the R software is a way to create one dimension FBM series  $B_H(t)$  [25]. Covariance of the increments of fractional Brownian motion is:

$$\operatorname{Cov}(u, v) = \frac{1}{2} \{ |u - (v+1)|_{2H} + |v - (u+1)|_{2H} - 2|u - v|_{2H} \}$$
(6)

Incorrect estimators (4) and (5) and incorrect classical hypothesis test statistics (2) were used to build the sequential monitoring processes without considering the covariance terms.

In the Equation (1),  $p_1$ ,  $p_2$  are the probability of success respectively in the Bernoulli distribution when the covariates  $X_1$ ,  $X_2$  are binary variables.  $\mu_1$ ,  $\mu_2$ ,  $\beta_1$ ,  $\beta_2$ ,  $p_1$ ,  $p_2$  were set up as 0.5, 0.5, 1, 1, 0.5, 0.5, respectively. 1000 replications were used for all the simulations. Patients were assumed to be sequentially randomized into two treatment groups by the block randomization (BR) (by blockrand() function in R software), stratified permuted block randomization (SPB) [26], and Pocock and Simon minimization designs (PS) [27] in the simulation studies consecutively. No covariate, two discrete covariates, and two continuous covariates situations were illustrated under misspecification scenarios in the simulation studies.

Assume 4000 patients were recruited in a clinical trial study with uniformly distributed enter time. An interim analysis would be done after every 100 new patients had finished the study. In total, 40 interim results were obtained. The maximum likelihood method was used to estimate the Hurst exponent (*H*) for normalized  $B_H$  value from the sequential test (3) in the entire paper. When H= 0.5, this indicated an uncorrelated process, corresponding to classical Brownian motion. It was shown that  $B_H$  has a long range dependence property when 0.5 < H < 1 [28,29].

The mean and standard deviation of Hurst exponent estimations were tabulated in Table 1, in which another Hurst estimation method proposed by Peltler Lévy Véhel was used to validate the MLE results [30]. Two Hurst estimation methods reached similar results. The distribution of the estimates of H is close to normal distribution. The visual histograms are shown in Figures 1–3. Mean estimated H values from the misspecification scenarios in Table 1 and Figures 1–3 deviated from 0.5. All test of statistical significance test proved this conclusion with p value less than 0.0001 by *t*-test (Student's *t*-Test) function in R software. According to the theoretical derivation results and simulation results, sequential test statistics do not follow a Brownian motion in the covariate adaptive randomized clinical trials sequential monitoring processes when error terms are correlated. Models with different covariate types reached similar conclusions.

		]	MLE Estimatio	n	Peltler Lévy Véhel H <sub>n,1</sub> (1)			
Designs		Mean	SD	p Value	Mean	SD	<i>p</i> Value	
	No covariate	0.5921	0.0287	< 0.0001	0.6056	0.0360	< 0.0001	
BR	Two discrete	0.5898	0.0288	< 0.0001	0.6077	0.0347	< 0.0001	
	Two continuous	0.5915	0.0292	< 0.0001	0.6067	0.0358	< 0.0001	
	No covariate	0.5504	0.0299	< 0.0001	0.5608	0.0347	< 0.0001	
PS	Two discrete	0.5503	0.0299	< 0.0001	0.5607	0.0348	< 0.0001	
	two continuous	0.5490	0.0300	< 0.0001	0.5604	0.0351	< 0.0001	
SPB	No covariate	0.5568	0.0302	< 0.0001	0.5685	0.0361	< 0.0001	
	Two discrete	0.5567	0.0302	< 0.0001	0.5694	0.0354	< 0.0001	
	Two continuous	0.5576	0.0286	< 0.0001	0.5691	0.0340	< 0.0001	

**Table 1.** Hurst exponent estimations for covariate adaptive randomized clinical trials,  $\varepsilon$ 's ~increments of FBM,  $\mu_1$ ,  $\mu_2$ ,  $\beta_1$ ,  $\beta_2$ ,  $p_1$ ,  $p_2$  (0.5, 0.5, 1, 1, 0.5, 0.5).



(a)





**Figure 1.** Histograms for estimated H values (BR design, *ε's~* increments of FBM, MLE method): (**a**) no covariate, (**b**) two discrete covariates, (**c**) two continuous covariates.



**Figure 2.** Histograms for estimated H values (PS design, *ε's~* increments of FBM, MLE method): (**a**) no covariate, (**b**) two discrete covariates, (**c**) two continuous covariates.



**Figure 3.** Histograms for estimated H values (SPB design, *ε's*~ increments of FBM, MLE method): (**a**) no covariate, (**b**) two discrete covariates, (**c**) two continuous covariates.

# 4. Conditional Power for Covariate Adaptive Randomized Clinical Trials with Correlated Error Structures

Test statistics derived from covariate adaptive randomized clinical trials do not follow a Brownian motion in the sequential monitoring processes when ignoring the error term covariance patterns in the statistics test formula [15]. The independent increment property is not satisfied under the misspecification assumption for the covariate adaptive randomized clinical trials. Therefore, a more general form of stochastic process, fractional Brownian motion, was proposed for the misspecification scenarios. In this section, conditional powers were calculated and compared between the BM and the FBM methods.

Conditional power is the probability of rejecting the null hypothesis, given the interim data results up to the time of interim analysis [31]. Conditional power is used for the decisions regarding possible early termination, or proceeding to the original plan, or increasing the sample sizes in the clinical trial interim analyses. The evaluation of the conditional power can predict the long-term clinical trial study results based on the partially observed data.

Under the null hypothesis of no treatment effect for the clinical trial,  $Z_t$  is the sequential test statistic at time t. Normalized  $Z_t$  from the estimated of  $\beta(t)$  in the covariate adaptive randomized clinical trial can be transformed to the B value as  $B_H(t) = Z_t t^{1/2}$ . Assuming time t > s, the conditional probability can be calculated with Brownian motion:  $CP = P(Z_t \ge a | Z_s = b) = 1 - \Phi(\frac{at^{1/2} - bs^{1/2}}{\sqrt{t-s}})$  [10,32]. When the treatment effect is not equal in two groups, Brownian motion would have a drift parameter  $\theta$ . Brownian motion with drift follows the normal distribution with mean  $\theta(t - s)$  and variance (t - s). The conditional power with drift is expressed as:  $CP = 1 - \Phi(\frac{at^{1/2} - bs^{1/2} - \theta(t-s)}{\sqrt{t-s}})$  [10,33].

The conditional probability under null hypothesis in terms of FBM was proposed in Lai et al. [34] expressed as:

$$CP = P(B_H(t_N) \ge s_N | B_H(t_n) = s_n \dots B_H(t_2) = s_2, \ B_H(t_1) = s_1) = 1 - \Phi\left(\frac{s_N - \mu_{H_n}}{\sigma_{H_n}}\right)$$
(7)

The conditional distribution in the Equation (7) is normally distributed as  $N(\mu_{H_n}, \sigma_{H_n}^2)$ .  $\Sigma$  is the variance- covariance matrix of  $B_H(t_N)$  and  $B_H(t_n) \dots B_H(t_2)$ ,  $B_H(t_1)$ , denoted as [ $\sum_{21}^{11} \sum_{22}^{12}$ ].  $\Sigma$  is a symmetric matrix.  $\mu_{H_n} = \sum_{12} \sum_{22}^{-1} \mu_n$ .  $\sigma_{H_n}^2 = \sum_{11} - \sum_{12} \sum_{22}^{-1} \sum_{21} \mu_n$  $= (B_H(t_n) \dots B_H(t_2), B_H(t_1))^t$  [34].

A two-stage simulation strategy was used to calculate the conditional power [35]. Recruited time *t* was divided into two parts by the time point *ti.fix*, which resembled the interim analysis time point in the real clinical trial. The first part before the time point *ti.fix*, which is called the fixed part, includes the already known data information. The second part after the time point *ti.fix* and until t = 1 is called the unknown part. The empirical result was denoted as the analyzed data from *ti.fix* to t = 1 by 1000 simulations. It was assumed that the empirical results can respond the consequence at the end of the whole clinical trial.

Based on the Formula (7), CPs with the assumptions of the BM and FBM were calculated respectively for the fixed part, from t = 0 to t = ti.fix. In Formula (7),  $s_1, s_2,..., s_n$  are the observed values before  $t_N$ . The mean  $\mu_{H_n}$  and  $\sigma_{H_n}$  are described in the definition of Equation (7). The conditional power was denoted as CP (BM) under BM assumption with H = 0.5. If the conditional power was calculated under the FBM assumption with the estimated H value by MLE approach, this was denoted as CP (FBM).

CP (empirical) was assumed to predict the results at the end of the whole clinical trial (t = 1) with the formula CP (empirical)= Count ( $Z(t_N) > 1.96$ )/1000. 1.96 is the critical value for the hypothesis test with alpha = 0.05 at t = 1. CP (empirical), among 1000 replications, was treated as the standard criterion to compare with theoretical values CP (BM) and CP (FBM) [35]. Scatter plots for CP (BM), CP (FBM) and empirical CPs virtually demonstrated the consistency between the theoretical and empirical data.

Scenarios were simulated when a Brownian motion showed a mean shift upwards with drift. Assume that B(t) is a standard Brownian motion process. A new Brownian motion process U(t) with drift would be shown as  $U(t) = \theta t + \sigma^2 B(t)$ ,  $t \ge 0$ .  $\theta$  is the drift parameter. Assuming  $\sigma^2 = 1$ , different parameters ( $\theta = 0$ , 0.5, 1, 1.5, 2) were illustrated in the conditional power simulation processes.  $\mu_{H_n}$  in the equation (7) corresponded with the drift parameters in the simulations. In addition, different interim time points, ti.fix = 0.7, 0.8, 0.9, were evaluated respectively.

In this section, a total sample size of 4000 participants was illustrated with 40 interim time points, although the actual number of interim analyses is much fewer. The error terms of the model were not independent and identically distributed but were assumed to follow increments of fractional Brownian motion with the Hurst exponent H = 0.9. In this study,  $B_H(t)$  were calculated first from sequential monitoring test statistics (3) under misspecification scenarios. H values were calculated based on the  $B_H(t)$ . Then, conditional powers were performed using Equation (7) with  $B_H(t)$  from sequential processes.

We evaluated the performances of CP (BM) and CP (FBM), compared with CP (empirical) under different drift parameters and *ti.fix* time points in Tables 2–8. Figures 4–9 were the scatter plots with x axis representing the theoretical CPs and y axis representing the empirical CPs. Green dots in the figures are the CP (FBM), while red dots are the CP (BM). Based on the results from the Tables, the mean estimations of the Hurst coefficients deviated away from 0.5 under different scenarios. Assuming the treatment effects are equal between two groups with drift parameter  $\theta = 0$ , CPs are small and close to 0. CPs increased dramatically along with the increase in drift parameter values. Different *ti.fix* = 0.7, 0.8, 0.9 resume the interim analysis time points in the simulations. CPs increased as *ti.fix* time values increased, since the interim analysis time point is closer to the end of the trial. When H > 0.5, most of the CP (FBM) values were greater than CP (BM) values. From all Tables, the majority of CP (FBM) values matched the CP (empirical) values better than the CP (BM) values.

$\mu_1, \mu_2, \beta_1, \beta_2, p_1, p_2$	ti.fix = 0.7	Designs	CP (Empirical)	CP (BM)	CP (FBM)	Н
		BR	0.0059	0.0062	0.0059	0.5890
(0.5, 0.5)	No covariate	PS	0.0094	0.0093	0.0093	0.5494
		SPB	0.0093	0.0095	0.0093	0.5569
	Two discrete	BR	0.0043	0.0049	0.0041	0.5875
(0.5, 0.5, 1, 1, 0.5, 0.5)		PS	0.0094	0.0093	0.0093	0.5494
· · · · · · ,		SPB	0.0123	0.0120	0.0123	0.5550
(0.5, 0.5, 1, 1)	Two continuous	BR	0.0050	0.0053	0.0050	0.5883
		PS	0.0139	0.0137	0.0141	0.5477
		SPB	0.0079	0.0081	0.0078	0.5583

**Table 2.** CPs under misspecification scenarios:  $\varepsilon's$ ~increments of FBM, H = 0.9, drift = 0.

**Table 3.** CPs under misspecification scenarios:  $\varepsilon$ 's~increments of FBM, H = 0.9, drift = 0.5.

$\mu_1, \mu_2, \beta_1, \beta_2, p_1, p_2$	ti.fix = 0.7	Designs	CP (Empirical)	CP (BM)	CP (FBM)	Н
		BR	0.0816	0.0776	0.0817	0.5889
(0.5, 0.5)	No covariate	PS	0.0931	0.0903	0.0926	0.5494
		SPB	0.0940	0.0903	0.0935	0.5568
	Two discrete	BR	0.0824	0.0774	0.0812	0.5874
(0.5, 0.5, 1, 1, 0.5, 0.5)		PS	0.0931	0.0903	0.0926	0.5493
		SPB	0.1020	0.0987	0.1019	0.5549
(0.5, 0.5, 1, 1)	Two continuous	BR	0.0760	0.0728	0.0758	0.5882
		PS	0.1013	0.0982	0.1010	0.5476
		SPB	0.0911	0.0879	0.0905	0.5582

$\mu_1, \mu_2, \beta_1, \beta_2, p_1, p_2$	ti.fix = 0.7	Designs	CP (Empirical)	CP (BM)	CP (FBM)	Н
(0.5, 0.5)	No covariate	BR PS SPB	0.3896 0.3861 0.3868	0.3647 0.3720 0.3704	0.3893 0.3852 0.3864	0.5887 0.5491 0.5566
(0.5, 0.5, 1, 1, 0.5, 0.5)	Two discrete	BR PS SPB	0.3927 0.3861 0.4011	0.3666 0.3720 0.3854	0.3920 0.3851 0.3996	0.5872 0.5490 0.5547
(0.5, 0.5, 1, 1)	Two continuous	BR PS SPB	0.3911 0.3831 0.3854	0.3630 0.3712 0.3700	0.3898 0.3820 0.3855	0.5880 0.5474 0.5580

**Table 4.** CPs under misspecification scenarios:  $\varepsilon$ 's~increments of FBM, H = 0.9, drift = 1.

**Table 5.** CPs under misspecification scenarios:  $\varepsilon$ '*s*~increments of FBM, H = 0.9, drift = 2.

$\mu_1, \mu_2, \beta_1, \beta_2, p_1, p_2$	ti.fix = 0.7	Designs	CP (Empirical)	CP (BM)	CP (FBM)	H
		BR	0.9788	0.9685	0.9784	0.5878
(0.5, 0.5)	No covariate	PS	0.9587	0.9528	0.9586	0.5482
		SPB	0.9618	0.9551	0.9619	0.5557
	Two discrete	BR	0.9779	0.9670	0.9774	0.5863
(0.5, 0.5, 1, 1, 0.5, 0.5)		PS	0.9587	0.9528	0.9586	0.5481
· · · · · · /		SPB	0.9634	0.9565	0.9629	0.5537
(0.5, 0.5, 1, 1)	Two continuous	BR	0.9771	0.9670	0.9767	0.5870
		PS	0.9570	0.9505	0.9567	0.5465
		SPB	0.9667	0.9594	0.9665	0.5571

**Table 6.** CPs under misspecification scenarios:  $\varepsilon$ 's~increments of FBM, H = 0.9, drift = 1.5, ti.fix = 0.7.

$\mu_1, \mu_2, \beta_1, \beta_2, p_1, p_2$	ti.fix = 0.7	Designs	CP (Empirical)	CP (BM)	CP (FBM)	Н
		BR	0.8100	0.7744	0.8081	0.5883
(0.5, 0.5)	No covariate	PS	0.7675	0.7507	0.7669	0.5487
		SPB	0.7732	0.7533	0.7727	0.5562
	Two discrete	BR	0.8010	0.7686	0.7997	0.5868
(0.5, 0.5, 1, 1, 0.5, 0.5)		PS	0.7675	0.7507	0.7669	0.5487
		SPB	0.7809	0.7616	0.7800	0.5543
(0.5, 0.5, 1, 1)	Two continuous	BR	0.8111	0.7747	0.8089	0.5876
		PS	0.7619	0.7451	0.7607	0.5470
		SPB	0.7766	0.7569	0.7760	0.5576

**Table 7.** CPs under misspecification scenarios:  $\varepsilon$ '*s*~increments of FBM, H = 0.9, drift = 1.5, *ti.fix* = 0.8.

$\mu_1, \mu_2, \beta_1, \beta_2, p_1, p_2$	<i>ti.fix</i> = 0.8	Designs	CP (Empirical)	CP (BM)	CP (FBM)	Н
(0.5, 0.5)	No covariate	BR PS	0.8705 0.8234	0.8447	0.8697 0.8231	0.5896 0.5491
		SPB	0.8360	0.8214	0.8356	0.5568
(0.5, 0.5, 1, 1, 0.5, 0.5)	Two discrete	BR PS	0.8571 0.8234	0.8348 0.8120	0.8568 0.8231	0.5877 0.5490
		SPB	0.8401	0.8260	0.8402	0.5558
(0.5, 0.5, 1, 1)	Two continuous	BR PS SPB	0.8701 0.8186 0.8374	0.8445 0.8071 0.8224	0.8695 0.8178 0.8372	0.5889 0.5476 0.5577

	$\mu_1, \mu_2, \beta_1, \beta_2, p_1, p_2$	ti.fix = 0.9	Designs	CP (Empirical)	CP (BM)	CP (FBM)	Н	
			BR	0.9041	0.8933	0.9041	0.5910	
	(0.5, 0.5)	No covariate	PS	0.8732	0.8658	0.8725	0.5496	
		SPB	0.8826	0.8746	0.8822	0.5576		
		BR	0.9052	0.8931	0.9049	0.5888		
	(0.5, 0.5, 1, 1, 0.5, 0.5)	Two discrete	PS	0.8732	0.8658	0.8725	0.5495	
		SPB	0.8857	0.8781	0.8856	0.5563		
(0.5, 0.5, 1, 1)		BR	0.9107	0.9002	0.9108	0.5905		
	(0.5, 0.5, 1, 1)	Two continuous	PS	0.8670	0.8609	0.8669	0.5482	
			SPB	0.8816	0.8739	0.8813	0.5574	

**Table 8.** CPs under misspecification scenarios:  $\varepsilon$ 's~increments of FBM, H = 0.9, drift = 1.5, *ti.fix* = 0.9.



**Figure 4.** Scatter plots for conditional power under misspecification scenarios (BR design) showing two continuous covariates with different *ti.fix* time points: (a) *ti.fix* = 0.7; (b) *ti.fix* = 0.8; (c) *ti.fix* = 0.9.



**Figure 5.** Scatter plots for conditional power under misspecification scenarios (BR design) two continuous covariates with different drift parameters  $\theta$ : (**a**)  $\theta = 0$ ; (**b**)  $\theta = 0.5$ ; (**c**)  $\theta = 1$ ; (**d**)  $\theta = 1.5$ ; (**e**)  $\theta = 2$ .



**Figure 6.** Scatter plots for conditional power under misspecification scenarios (PS design) showing no covariates with different *ti.fix* time points: (**a**) *ti.fix* = 0.7; (**b**) *ti.fix* = 0.8; (**c**) *ti.fix* = 0.9.



**Figure 7.** Scatter plots for conditional power under misspecification scenarios (PS design) showing no covariates with different drift parameters  $\theta$ : (**a**)  $\theta$  = 0; (**b**)  $\theta$  = 0.5; (**c**)  $\theta$  = 1; (**d**)  $\theta$  = 1.5; (**e**)  $\theta$  = 2.



**Figure 8.** Scatter plots for conditional power under misspecification scenarios (SPB design) showing two discrete covariates with different *ti.fix* time points: (a) *ti.fix* = 0.7; (b) *ti.fix* = 0.8; (c) *ti.fix* = 0.9.



**Figure 9.** Scatter plots for conditional power under misspecification scenarios (SPB design) showing two discrete covariates with different drift parameters  $\theta$ : (a)  $\theta = 0$ ; (b)  $\theta = 0.5$ ; (c)  $\theta = 1$ ; (d)  $\theta = 1.5$ ; (e)  $\theta = 2$ .

# 5. Conclusions and Discussions

In this study, we investigated the sequential monitoring properties in covariate adaptive randomized clinical trials under the misspecification scenarios. We also performed numerical simulations under various situations in which the mean estimates of Hurst coefficient by MLE from the sequential test statistics under misspecification scenarios deviated from 0.5. Brownian motion is satisfied only when H = 0.5. Therefore, the independent increment assumption was violated and Brownian motion was not appropriate for the sequential process. However, clinical researchers may not know the existence of the covariance in the error terms, and hence use the original classical statistic test under the misspecification scenarios, leading to non-Brownian motion trajectory of the test statistics under sequential analysis. Therefore, it is necessary to estimate the Hurst coefficient.

We calculated conditional powers for covariate adaptive randomized clinical trials with mis-specified error structures of the model under different covariate types, adaptive designs, drift parameters, and interim time points. Conditional powers based on the fractional Brownian motion (CP (FBM)) assumption resulted in better consistency with the standard empirical value (CP (empirical)) than conditional powers under the classical Brownian motion (CP (BM)) assumption. When the  $H_{MLE} > 0.5$ , most conditional powers under the FBM assumption were greater than the conditional powers under the classical Brownian motion assumption. The fractional Brownian motion, incorporating a dependent increment assumption, would be a reasonable approach for the clinical trial sequential analyses. Even if the sequential procedure actually follows the Brownian motion, the application of the fractional Brownian motion technique would still be useful, since BM is a special case of FBM with H = 0.5.

**Author Contributions:** Conceptualization, D.L. and H.Z.; methodology, D.L. and H.Z.; software, Y.Y.; validation, D.L., H.Z. and Y.Y.; formal analysis, Y.Y.; investigation, D.L., H.Z. and Y.Y.; resources, D.L.; data curation, Y.Y.; writing—original draft preparation, Y.Y.; writing—review and editing, D.L. and H.Z.; visualization, Y.Y.; supervision, D.L.; project administration, D.L.; funding acquisition, D.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was partially supported by Cancer Prevention Research Institute of Texas (RP170668).

**Acknowledgments:** This work was partially supported by Cancer Prevention Research Institute of Texas (RP170668).

Conflicts of Interest: The authors declare no conflict of interest.

### References

- 1. Chow, S.-C.; Liu, J.-P. Design and Analysis of Clinical Trials: Concepts and Methodologies; John Wiley & Sons: Hoboken, NJ, USA, 2008; Volume 507.
- Pallmann, P.; Bedding, A.W.; Choodari-Oskooei, B.; Dimairo, M.; Flight, L.; Hampson, L.V.; Holmes, J.; Mander, A.P.; Odondi, L.; Sydes, M.R.; et al. Adaptive designs in clinical trials: Why use them, and how to run and report them. *BMC Med.* 2018, 16, 1–15. [CrossRef]
- 3. Hu, Y.Q.; Hu, F.F. Asymptotic Properties of Covariate-Adaptive Randomization. Ann. Stat. 2012, 40, 1794–1815. [CrossRef]
- Khan, O.; Fotheringham, S.; Wood, V.; Stimson, L.; Zhang, C.L.; Pezzella, F.; Duvic, M.; Kerr, D.J.; La Thangue, N.B. HR23B is a biomarker for tumor sensitivity to HDAC inhibitor-based therapy. *Pro. Natl. Acad. Sci. USA* 2010, 107, 6532–6537. [CrossRef] [PubMed]
- 5. Yu, J.; Lai, D.J. Sequential monitoring of covariate adaptive randomized clinical trials with sample size re-estimation. *Contemp Clin. Trials* **2019**, *87*, 105874. [CrossRef]
- 6. Efron, B. Forcing a sequential experiment to be balanced. *Biometrika* 1971, 58, 403–417. [CrossRef]
- Ma, W.; Hu, F.F.; Zhang, L.X. Testing Hypotheses of Covariate-Adaptive Randomized Clinical Trials. J. Am. Stat. Assoc. 2015, 110, 669–680. [CrossRef]
- 8. Bugni, F.A.; Canay, I.A.; Shaikh, A.M. Inference Under Covariate-Adaptive Randomization. J. Am. Stat. Assoc. 2018, 113, 1784–1796. [CrossRef]
- 9. Ma, W.; Qin, Y.C.; Li, Y.; Hu, F.F. Statistical Inference for Covariate-Adaptive Randomization Procedures. J. Am. Stat. Assoc. 2020, 115, 1488–1497. [CrossRef]
- 10. Lan, K.K.; Wittes, J. The B-value: A tool for monitoring data. Biometrics 1988, 44, 579-585. [CrossRef]
- 11. Proschan, M.A.; Lan, K.G.; Wittes, J.T. *Statistical Monitoring of Clinical Trials: A Unified Approach*; Springer: New York, NY, USA, 2006.
- 12. Zhu, H.J.; Hu, F.F. Sequential Monitoring of Covariate-Adaptive Randomized Clinical Trials. Stat. Sin. 2019, 29, 265–282.

- 13. Abundo, M.; Pirozzi, E. On the Integral of the Fractional Brownian Motion and Some Pseudo-Fractional Gaussian Processes. *Mathematics* **2019**, *7*, 991. [CrossRef]
- 14. Wang, P.G.; Xu, Y. Averaging Method for Neutral Stochastic Delay Differential Equations Driven by Fractional Brownian Motion. *J. Funct. Space* **2020**, 2020, 5212690. [CrossRef]
- 15. Yang, Y. Clinical trial designs and monitoring with fractional Brownian motion. Proquest Diss. Publ. 2020.
- Ferguson, T.S.; Klass, M.J. A representation of independent increment processes without Gaussian components. *Ann. Math. Stat.* 1972, 43, 1634–1643. [CrossRef]
- 17. Hurst, H.E. Long-term storage capacity of reservoirs. Trans. Am. Soc. Civ. Eng. 1951, 116, 770–799. [CrossRef]
- 18. Taqqu, M.S. Fractional Brownian motion and long-range dependence. In *Theory and Applications of Long-Range Dependence;* Doukhan, P., Oppenheim, G., Taqqu, M.S., Eds.; Birkhäuser: Boston, MA, USA, 2003.
- 19. Yerlikaya-Ozkurt, F.; Vardar-Acar, C.; Yolcu-Okur, Y.; Weber, G.W. Estimation of the Hurst parameter for fractional Brownian motion using the CMARS method. *J. Comput. Appl. Math.* **2014**, *259*, 843–850. [CrossRef]
- 20. Mandelbrot, B.B.; Van Ness, J.W. Fractional Brownian motions, fractional noises and applications. *Siam Rev.* **1968**, *10*, 422–437. [CrossRef]
- 21. McCauley, J.L.; Gunaratne, G.H.; Bassler, K.E. Hurst exponents, Markov processes, and fractional Brownian motion. *Phys. A* 2007, 379, 1–9. [CrossRef]
- 22. Lai, D.J. Estimating the Hurst effect and its application in monitoring clinical trials. *Comput. Stat. Data* 2004, 45, 549–562. [CrossRef]
- 23. Jennison, C.; Turnbull, B.W. Group Sequential Methods with Applications to Clinical Trials; CRC Press: Boca Raton, FL, USA, 1999.
- 24. Qian, H. Fractional Brownian motion and fractional Gaussian noise. In *Processes with Long-Range Correlations;* Springer: Berlin/Heidelberg, Germany, 2003; pp. 22–33.
- 25. Huang, J. Somebm: Some Brownian Motions Simulation Functions. R Package Version 0.1. Available online: https://CRAN.R-project.org/package=somebm (accessed on 5 September 2021).
- 26. Zelen, M. The randomization and stratification of patients to clinical trials. J. Chronic. Dis. 1974, 27, 365–375. [CrossRef]
- 27. Pocock, S.J.; Simon, R. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics* **1975**, *31*, 103–115. [CrossRef]
- 28. Ortigueira, M.D.; Batista, A.G. A fractional linear system view of the fractional Brownian motion. *Nonlinear Dynam.* **2004**, *38*, 295–303. [CrossRef]
- 29. Beran, J. Statistical methods for data with long range dependence. Stat. Sci. 1992, 7, 4404–4427.
- 30. Peltier, R.F.; Lévy Véhel, J. A new method for estimating the parameter. *Research Report*. 1994. RR-2396, inria (inria-00074279). Available online: https://hal.inria.fr/inria-00074279 (accessed on 5 September 2021).
- 31. Lachin, J.M. A review of methods for futility stopping based on conditional power. *Stat. Med.* **2005**, *24*, 2747–2764. [CrossRef] [PubMed]
- Lai, D.J.; Moye, L.A.; Davis, B.R.; Brown, L.E.; Sacks, F.M. Brownian motion and long-term clinical trial recruitment. J. Stat. Plan Infer. 2001, 93, 239–246. [CrossRef]
- 33. Moyé, L.A. Statistical Monitoring of Clinical Trials: Fundamentals for Investigators; Springer: New York, NY, USA, 2006.
- 34. Lai, D.J.; Davis, B.R.; Hardy, R.J. Fractional Brownian motion and clinical trials. J. Appl. Stat. 2000, 27, 103–108. [CrossRef]
- 35. Huang, V.J.; Lai, D.J. Interim Analysis of Clinical Trials: Simulation Studies of Conditional Power under Fractional Brownian Motion. *Fractals* **2019**, *27*, 1950133. [CrossRef]