Statistica Sinica **9**(1999), 939-952

BIAS CORRECTION IN GROUP SEQUENTIAL ANALYSIS WITH CORRELATED DATA

Roger P. Qu and David L. DeMets

R. W. Johnson Pharmaceutical Research Institute and University of Wisconsin-Madison

Abstract: This paper focuses on the bias of the group sequential estimate of treatment effect for correlated data using the generalized estimating equation (GEE) method and the Lan and DeMets alpha-spending function. Linear and logistic regressions are used to examine (a) the magnitude of the bias of a sequential estimate with correlated data; (b) the influence of the true correlation structure on bias. A bias-corrected sequential estimate is proposed using a Brownian motion approximation and numerical simulation. Logistic regression is used to illustrate and to assess the performance of the proposed method.

Key words and phrases: Alpha-spending function, Brownian motion, correlation structure, GEE, interim analysis, linear regression, logistic regression, sequential boundaries, simulation, working correlation.

1. Introduction

Introduced by Pocock (1977) and O'Brien-Fleming (1979), group sequential analysis has increasingly become a standard method in clinical trials. It has been known that sequential analysis is prone to exaggeration of treatment effect due to early stopping of a trial (Emerson and Fleming (1990), Hughes and Pocock (1988), Hughes, Freedman and Pocock (1992), Pinheiro and DeMets (1997) and Li and DeMets (1996)). The boundaries in group sequential methods are determined under the null hypothesis for the protection of the overall type-I error rate. When the null hypothesis is not true, the trial will stop only when the sequential estimate is large in absolute value, which leads to a biased estimate of the treatment effect. Traditional estimates of treatment effect are not always adjusted for sequential monitoring.

Whitehead (1986) proposed an adjustment to the maximum likelihood estimate following a sequential probability ratio or triangular test. Whitehead's method relies on the feasibility of an assessment of the sequential bias when the treatment effect is given. With a known correlation structure for correlated data, the distribution of the sequential estimate can be approximated by Brownian motion, and thus its bias can be assessed (Pinheiro and DeMets (1997)). When the true correlation structure of the observations is unknown, the performance of bias reduction using Brownian motion is unclear. Since the GEE method for correlated data (Liang and Zeger (1986)) does not require knowledge of the true correlation structure given the alpha-spending function, the boundaries can be determined using the asymptotic normality of the GEE estimate and Mulnor's subroutine (Schervish (1984)) for calculation of multivariate normal boundaries (Wei, Su and Lachin (1990) and Lee and DeMets (1991)). But the lack of the known correlation structure prevents one from assessing the bias of the sequential estimate, and thus Whitehead's adjustment does not seem feasible. Perhaps the most forbidding factor in using Whitehead's adjustment to the GEE estimate is the extremely low speed of Mulnor's subroutine when the dimension of the multivariate normal distribution is higher than 3.

The paper is organized as follows. First, group sequential analysis for correlated data is briefly reviewed in Section 2. In Section 3, bias of the group sequential estimate using the independence working correlation in GEE estimation, and the alpha-spending function in conjunction with the multivariate integration method, is examined for linear and logistic regressions. The influence of the true correlation structure on the sequential bias is studied using one-dependent, AR-1, exchangeable and independence structures. A bias-corrected sequential estimate using the independence working correlation structure in the GEE estimation and Brownian motion approximation is proposed in Section 4. The procedure is illustrated using logistic regression. Finally, Section 5 contains some discussion.

2. Group Sequential Analysis for Correlated Data

Let K be the total number of interim analyses, and $N(k)$ be the number of subjects available for the kth interim analysis, $k = 1, \ldots, K$. Let $\boldsymbol{\beta}$ be the unknown parameters, γ the component of $\boldsymbol{\beta}$ corresponding to the treatment effect of interest, and $\hat{\gamma}^{(k)}$ the estimate of γ at the kth interim analysis using the GEE method. Then, under regularity conditions, the joint distribution of $(\hat{\gamma}^{(1)}, \ldots, \hat{\gamma}^{(k)})$ is asymptotically normal with mean $(\gamma, \ldots, \gamma)$. For a detailed description of GEE estimation for group sequential analysis, see Gange and DeMets (1996).

Let $\psi_k = \mathrm{Var}^{-1}(\hat{\gamma}^{(k)})$, $k = 1, \ldots, K$. Then the amount of information accumulated by the kth interim analysis is defined as ψ_k, where ψ_K measures the total information available by the end of the trial. The information fraction at the kth analysis is $t_k^* = \psi_k/\psi_K = \mathrm{Var}\,(\hat{\gamma}^{(K)})/\mathrm{Var}\,(\hat{\gamma}^{(k)})$.

The information fraction t^* is needed for determination of boundaries for sequential monitorings. Once t_k^* is determined, or estimated using calendar time or number of responses observed or other means, one can use the alpha-spending function $\alpha(\cdot)$ of Lan and DeMets (1983) to calculate the amount of significance level spent at the kth analysis. Given t_k^*, the sequential boundaries $c_1, \ldots, c_K$ are calculated via

$$\mathrm{P}\left(\left|\hat{\gamma}^{(1)}\sqrt{\psi_1}\right| \le c_1, \ldots, \left|\hat{\gamma}^{(k-1)}\sqrt{\psi_{k-1}}\right| \le c_{k-1}, \left|\hat{\gamma}^{(k)}\sqrt{\psi_k}\right| > c_k\right) = \alpha(t_k^*) - \alpha(t_{k-1}^*),$$

for $k = 1, \ldots, K$.

In the following, the sequential bias is examined for four true correlation structures: one-dependent $\boldsymbol{R}_1$, exchangeable $\boldsymbol{R}_2$, AR-1 $\boldsymbol{R}_3$ and independence $\boldsymbol{R}_4 = \boldsymbol{I}$ (Liang and Zeger (1986)). When the working correlation matrix $\boldsymbol{R}$ is the true correlation matrix, the sequence $\{\hat{\gamma}^{(j)}\}$ has an independent increment structure, i.e., for $j < k < l$, $\mathrm{Cov}\,(\hat{\gamma}^{(k)} - \hat{\gamma}^{(j)}, \hat{\gamma}^{(l)} - \hat{\gamma}^{(k)}) = 0$ (Gange and DeMets (1996)). Hence the Brownian motion approximation and existing software (Reboussin, DeMets, Kim and Lan (1992)) can be used to calculate the sequential boundaries. Even if $\boldsymbol{R}$ is not the true correlation matrix, the joint distribution of $(\hat{\gamma}^{(1)}, \ldots, \gamma^{(k)})$ is still asymptotically normal. Using the asymptotic normality of $(\hat{\gamma}^{(1)}, \ldots, \hat{\gamma}^{(k)})$ and Mulnor's subroutine, one can still calculate the asymptotic sequential boundaries using the alpha-spending function.

3. Bias of Group Sequential Estimate

We now examine the bias associated with group sequential analysis for correlated data. For independent data, the naive or MLE estimate following early stopping has considerable bias. Let $\hat{\gamma}$ denote the estimate of γ at stopping time τ, where τ is defined as $\inf\{k|\ |\hat{\gamma}^{(k)}\sqrt{\psi_k}| > c_k, k = 1, \ldots, K-1\}$, or $\tau = K$ otherwise. The expectation of $\hat{\gamma}$ equals

$$\begin{aligned}\mathrm{E}(\hat{\gamma}) = \Sigma_{k=1}^{K} \mathrm{E}&\left\{\hat{\gamma}^{(k)} \mathrm{I}_{[|\hat{\gamma}^{(j)}\sqrt{\psi_j}| \le c_j, j=1,\ldots,k-1, |\hat{\gamma}^{(k)}\sqrt{\psi_k}| > c_k]}\right\} \\ +\mathrm{E}&\left\{\hat{\gamma}^{(K)} \mathrm{I}_{[|\hat{\gamma}^{(j)}\sqrt{\psi_j}| \le c_j, j=1,\ldots,K]}\right\},\end{aligned}$$

where $\mathrm{I}_{[A]}$ is the indicator function of an event A. Let $b(\gamma)$ denote the bias of $\hat{\gamma}$, i.e., $b(\gamma) = \mathrm{E}(\hat{\gamma}) - \gamma$. One can see that if $\gamma = 0$, i.e., there is no treatment effect, then $b(0) = 0$ since the joint distribution of $\{\hat{\gamma}^{(i)}\}$ in this situation is symmetric about zero. If γ is positive, then for $K = 2$, the simplest case with only one interim monitoring,

$$\begin{aligned}b(\gamma) &= \mathrm{E}(\hat{\gamma}^{(1)}) + \mathrm{E}\left\{(\hat{\gamma}^{(2)} - \hat{\gamma}^{(1)}) \mathrm{I}_{[|\hat{\gamma}^{(1)}\sqrt{\psi_1}| \le c_1]}\right\} - \gamma \\ &\approx \mathrm{E}\left\{(\hat{\gamma}^{(2)} - \hat{\gamma}^{(1)}) \mathrm{I}_{[\hat{\gamma}^{(1)}\sqrt{\psi_1} \le c_1]}\right\}.\end{aligned}$$

Note that only when $\hat{\gamma}^{(1)}\sqrt{\psi_1}$ is small, is the difference $\hat{\gamma}^{(2)} - \hat{\gamma}^{(1)}$ taken into account in the bias calculation. Thus $\hat{\gamma}$ tends to be an over-estimate for positive

γ. Also, the later the first interim analysis, the closer $\hat{\gamma}^{(2)}$ and $\hat{\gamma}^{(1)}$, and thus the smaller the bias in general.

The marginal distributions we examine here are Gaussian and binomial. The former corresponds to the usual linear regression with normal deviates and the latter to the logistic regression. Assume that the observations are from a randomized clinical trial with treatment effect γ, and let y denote the response and x be 1 if the subject receives treatment, 0 otherwise. The pair-wise correlation ρ is 0.3 or 0.7, reflecting different levels of correlation. The sample size n is 100 with 50 subjects in each group. The true correlation structures used in the simulation are the one-dependent $\boldsymbol{R}_1$, exchangeable $\boldsymbol{R}_2$, AR-1 $\boldsymbol{R}_3$ and independence $\boldsymbol{R}_4$, while the working correlation is the independence structure. For simplicity, assume that observations from each subject are evenly spaced throughout the course of the trial. Thus the design of the trial is completely balanced in terms of data collection. The simulation was performed with 2000 replicates. The null hypothesis is $\gamma = 0$, and the alternative $\gamma \neq 0$ with the type-I error equal to 0.05 for a two-sided test.

3.1. Linear regression

For linear regression, $y = \xi + \gamma x + \epsilon$, $\epsilon \sim \mathrm{N}(0,\sigma^2)$. Assume $\sigma^2 = 1$ and $\xi = 1$. Suppose that there are ten observations from each subject and that the trial is monitored at the times when 2, 5, 7 and 10 observations from each subject are collected. Thus the number of subjects available at each interim analysis is $N(1) = N(2) = N(3) = N(4)$ using previous notation.

Figure 1 presents plots of bias for $\rho = 0.3$ (smoothed using a second degree local regression, similarly for Figures 2, 3, 4 and 5). Overall, the O'Brien-Fleming boundaries have smaller bias than Pocock's boundaries. The power is almost the same for the O'Brien-Fleming and Pocock boundaries, with their average power shown in Figure 1. Pocock's boundaries produce a large bias for small values of γ, while O'Brien-Fleming's boundaries have their peaks when Pocock's boundaries are almost unbiased. However, this occurs for an extremely large, almost unlikely treatment effect. The Pocock boundaries are smaller than the O'Brien-Fleming boundaries for early monitoring and thus produce larger bias than O'Brien-Fleming boundaries for a small treatment effect. For the O'Brien-Fleming boundaries, since it is very hard to stop early for a small treatment effect, the trial will be carried almost to its designed end and thus the bias is much smaller than for the Pocock boundaries. These properties are consistent with those observed by others such as Pinheiro and DeMets (1997) and Li and DeMets (1996).

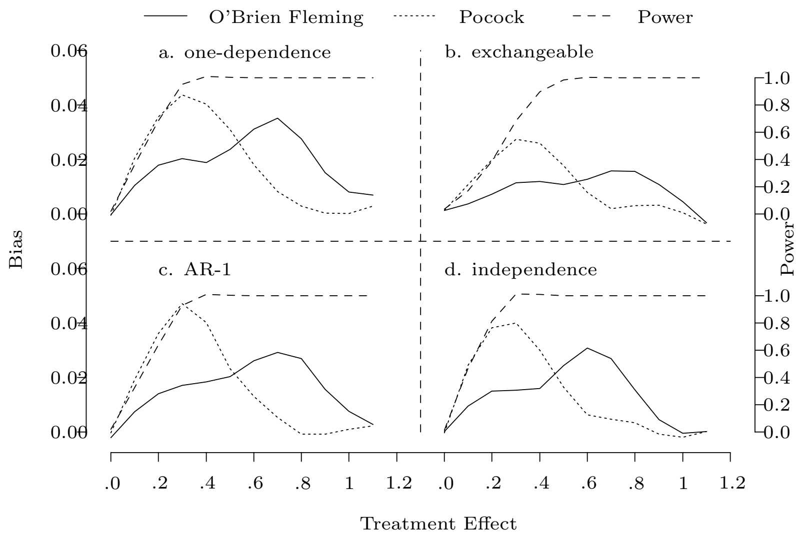

Figure 1. Bias and average power for linear regression with $\rho = 0.3$.

The influence of true correlation structure is clearly demonstrated here by the similar bias patterns of plots a, c, and d of Figure 1. For the one-dependent structure, only adjacent observations are correlated. For the AR-1 structure, correlation of observations more than one-lag apart is almost negligible. Thus biases from these correlation structures behave almost the same as for the independence correlation structure. Interestingly, the exchangeable structure has much smaller bias compared to other structures, including even the independence correlation. In fact, because of the correlation among all observations from a single subject, $(\hat{\gamma}^{(1)}, \ldots, \hat{\gamma}^{(k)})$ are more correlated with each other than for the other three correlation structures, and hence produce less bias. The value of pair-wise correlation ρ has little influence on the overall level of correlation for the three correlation structures $\boldsymbol{R}_1$, $\boldsymbol{R}_2$, and $\boldsymbol{R}_4$, and thus their bias pattern and magnitude for $\rho = 0.7$ are almost the same as for $\rho = 0.3$. Figure 2 shows the bias for the exchangeable structure with $\rho = 0.7$. For the exchangeable structure, the magnitude of bias for $\rho = 0.7$ should be less than for $\rho = 0.3$ since interim estimates are more correlated for larger values of ρ than for smaller values.

3.2. Logistic regression

Suppose that the response y is binary with possible values 0 and 1, and let p be the probability that y is equal to 1. Thus $p = e^{\xi+\gamma x}/(1 + e^{\xi+\gamma x})$. Assume that $\xi = .1$. Suppose there are eight observations from each subject and that the trial is monitored at the times when 2, 4, 6 and 8 observations from each subject are collected. Figure 3 is the counterpart of Figure 1 for logistic regression. Again, bias for one-dependent, AR-1 and independence structures is almost the same, and the exchangeable structure has smaller bias than the other three. For Pocock's boundaries, the pattern of bias is similar to the linear regression case; for O'Brien-Fleming's boundaries, the bias keeps increasing for the range of γ examined. The probability of 1 is 0.52 and 0.83 for γ equal to 0 and 1.5 respectively. Thus the range of increase of response rate from control to treatment group is from 0% to 60% for $\gamma \in (0, 1.5)$. The average power reaches 100% when $\gamma = 1.0$ for all four situations. Thus O'Brien-Fleming's boundaries produce less bias than Pocock's for most practical situations. Similar to the linear regression case, the bias for $\rho = 0.7$ is almost the same as for $\rho = 0.3$ for the one-dependent, AR-1 and independence structures, and smaller than for $\rho = 0.3$ for the exchangeable structure (not shown here).

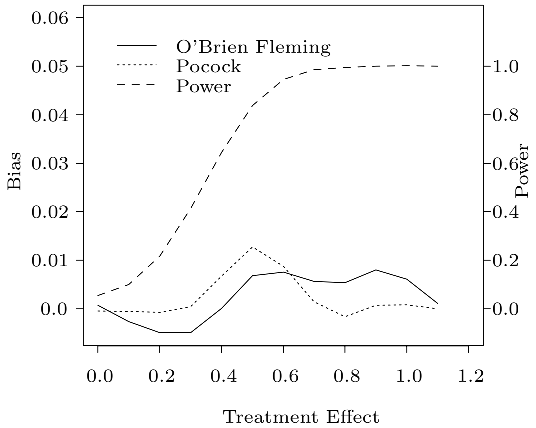

Figure 2. Bias and average power for linear regression for true exchangeable correlation structure with $\rho = 0.7$.

3.3. Influence of different monitoring patterns

Simulations comparing different monitoring schemes in various scenarios revealed that early monitoring schemes have larger bias than later schemes for both O'Brien-Fleming and Pocock's boundaries. See Qu and DeMets (1998) for details. Similar results were observed by others, including Li and DeMets (1996) and Pinheiro and DeMets (1997).

4. Bias Reduction of Group Sequential Estimate

4.1. Bias estimation

We now consider the Whitehead bias correction. The Whitehead bias-corrected estimate γ^* is defined as the solution of

$$\gamma^* = \hat{\gamma} - b(\gamma^*). \tag{1}$$

The bias of γ^* is $b^*(\gamma^*) = \mathrm{E}(\gamma^*) - \gamma = [\mathrm{E}(\hat{\gamma}) - \gamma] - \mathrm{E}[b(\gamma^*)] = b(\gamma) - \mathrm{E}[b(\gamma^*)]$. So the bias of $\hat{\gamma}$ is reduced by the amount of $\mathrm{E}[b(\gamma^*)]$. See Li and DeMets (1996) for more discussion of the properties of γ^*.

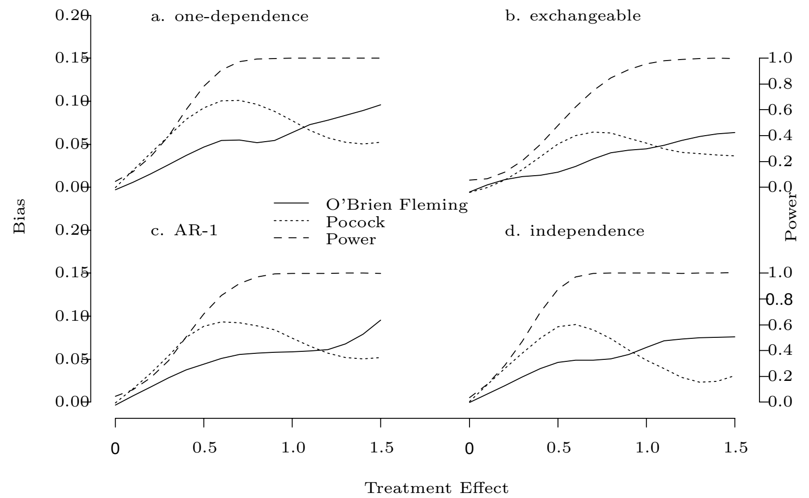

Figure 3. Bias and average power for logistic regression with $\rho = 0.3$.

In practice, (1) can be solved numerically using Newton-Raphson iteration

$$\gamma_i^* = \gamma_{i-1}^* + \frac{(\hat{\gamma} - \gamma_{i-1}^*) - b(\gamma_{i-1}^*)}{1 + b'(\gamma_{i-1}^*)}, i = 1, 2, \ldots,$$

where $b'(\gamma)$ denotes the derivative of the bias function $b(\gamma)$, and $\gamma_0^* = \hat{\gamma}$. The iteration requires evaluation of $b'(\gamma)$ which can be done numerically. The key in using the Whitehead correction is the estimate of the sequential bias given the treatment effect. In the following, we are going to use Brownian motion to approximate the sequential bias $b(\gamma)$.

Recall the information fraction $t_k^* = \psi_k/\psi_K$. If the working correlation is the true correlation, then for $k < l$, $\mathrm{Cov}\,(\hat{\gamma}^{(k)}, \hat{\gamma}^{(l)}) = \mathrm{Var}\,(\hat{\gamma}^{(l)})$ (Gange and DeMets (1996)). As a result

$$\begin{pmatrix} \hat{\gamma}^{(1)} \\ \ldots \\ \hat{\gamma}^{(K)} \end{pmatrix} \sim \mathrm{N}\left[\begin{pmatrix} \gamma \\ \ldots \\ \gamma \end{pmatrix}, \psi_K^{-1} \left(\frac{\min(t_i^*, t_j^*)}{t_i^* \cdot t_j^*} \right)_{i,j} \right]. \tag{2}$$

Define

$$B_N(t_k^*) = \sqrt{t_k^*}\, \hat{\gamma}^{(k)} \sqrt{\psi_k}.$$

Then, we have (a) $\mathrm{E}[B_N(t_k^*)] = t_k^* \cdot (\sqrt{\psi_K} \cdot \gamma)$; (b) $\mathrm{Var}\,[B_N(t_k^*)] = t_k^*$; and (c) $\mathrm{Cov}\,[B_N(t_k^*), B_N(t_l^*)] = \min\{t_k^*, t_l^*\}$. Hence, $B_N(t^*)$ resembles a Brownian motion on [0,1]. Thus, instead of Mulnor's subroutine, one can use the existing subroutine of Reboussin, DeMets, Kim and Lan (1992) for the calculation of sequential boundaries. Since (2) is the true joint distribution of the interim estimate, one can use it to estimate the bias of the sequential estimate of γ. Given the sequential boundaries $\{c_k\}$ calculated under null hypothesis $\gamma = 0$, and a value of γ, one proceeds as follows:

1. Calculate $\psi_1, \ldots, \psi_K$ and $t_1^*, \ldots, t_K^*$ under the alternative hypothesis;
2. Generate $\hat{\gamma}^{(1)}, \ldots, \hat{\gamma}^{(K)}$ from distribution (2);
3. Compare $\left\{\left|\hat{\gamma}^{(k)}\ \sqrt{\psi_k}\right|\right\}$ with $\{c_k\}$ and calculate $\hat{\gamma} = \hat{\gamma}^{(\tau)}$, where τ is the stopping time;
4. repeat steps 2 and 3 a total of M times.

Let $\{\hat{\gamma}_i\}_{i=1}^M$ be the M estimates of γ. The sequential bias of γ is then estimated as

$$b(\gamma) = \Sigma_{i=1}^M \frac{\hat{\gamma}_i}{M} - \gamma.$$

For the sake of simulations in this paper, $\{\psi_k\}$, $\{t_k^*\}$ and $\{c_k\}$ are calculated using the known values of the nuisance parameters involved in the models. In practice, one has to calculate ψ_k, t_k^* and c_k under the null hypothesis using the updated estimate of nuisance parameters at each monitoring. When the study is stopped at some interim monitoring with the sequential estimate $\hat{\gamma}$, one can follow the above four-step procedure to estimate the bias of the sequential estimate.

4.2. Numerical simulations

The above procedure is based on the assumption that the working correlation $\boldsymbol{R}$ is the true correlation structure. In practice, the true correlation is unknown, one has to choose a working correlation, and one of the possible choices is the independence working correlation. If the independence structure is close to the true correlation, then this procedure should give satisfactory results. If the independence structure is far from correct, the general performance of using the independence structure is unknown. For the previous logistic regression example with $\rho = 0.3$, Figure 4 shows the bias of the original and bias-corrected sequential estimate using the O'Brien-Fleming boundaries. The bias correction is surprisingly good even for the true exchangeable structure, and the amount of bias reduction is almost the same for all four structures. Since the bias for the exchangeable model is relatively smaller than that for the other three structures, the procedure over-corrects the sequential estimate slightly. But the magnitude of over-correction is almost negligible compared to the original bias. In fact, no matter what the true correlation structure is, the mean of (2) is always correct. Whether the working correlation is the true structure or not only influences the variance structure of (2).

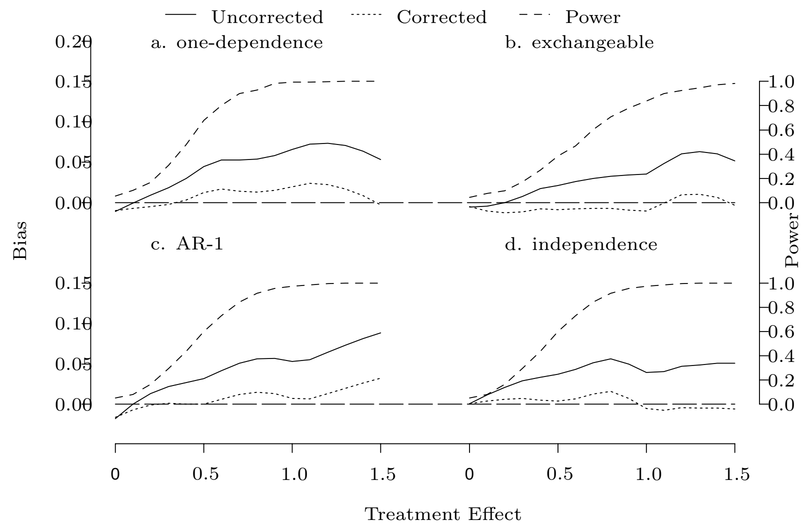

Figure 4. Bias reduction using O'Brien-Fleming's boundaries and the independence working correlation structure for logistic regression with $\rho = 0.3$.

The first half of Table 1 summarizes the operating characteristics of the original and bias-corrected sequential estimates for the true one-dependent correlation structure for $\gamma = 0, 1.4(0.2)$. Although the bias of the corrected estimate is much smaller than it is for the original estimate, its standard deviation is almost the same. This is consistent with the fact that the bias-corrected estimate actually shrinks the original sequential estimate. The ratio of bias to the true value of γ is significantly smaller for the corrected bias than for the original bias. The second half of Table 1 presents the results for the true exchangeable structure for which the standard deviation of the bias-corrected estimate is slightly smaller than that of the original. Since the true one-dependent structure is almost the same as the independence working correlation, its standard deviation is smaller than for the exchangeable case. The comparison of operating characteristics of the two estimates for the true AR-1 and independence structures is much the same as for the one-dependent and thus is not shown here.

The last column of Table 1 shows the difference between the bias of the initial estimate $b(\gamma)$ and that of the bias-corrected estimate $b^*(\gamma)$. For the values of γ in the table, one can see that the difference between these two biases is almost the same for the one-dependent and exchangeable structures. This is also true for other values of γ in our simulations and for all the four correlation structures examined in this paper in our simulations (not shown here). Once the initial estimate $\hat{\gamma}$ is obtained, the bias-corrected estimate is derived using (1) in conjunction with (2). Although the initial estimate is biased, its magnitude is mainly determined by the true value of γ. Thus the bias-corrected estimate is primarily determined by the true value of γ through distribution (2). If the initial estimate is very close to the true value, then the bias-corrected estimate will over-correct the bias. Nevertheless, the overall bias reduction using Whitehead adjustment and Brownian motion estimation is substantial compared to the original estimate.

Table 1. Operating characteristics of sequential and bias-corrected estimates using O'Brien-Fleming's boundaries for logistic regression with true one-dependent or exchangeable correlation structure with $\rho = 0.3$.

	γ	b(γ)	std($\hat{\gamma}$)	$b(\gamma)/\gamma$	$b^*(\gamma)$	$std(\gamma^*)$	$b^*(\gamma)/\gamma$	$b(\gamma) - b^*(\gamma)$
One-dependence								
	0.0	–0.015	0.209	*	–0.013	0.196	*	–0.001
	0.2	0.011	0.228	0.059	–0.002	0.214	–0.009	0.013
	0.4	0.024	0.247	0.062	–0.001	0.235	–0.004	0.027
	0.6	0.047	0.262	0.079	0.011	0.255	0.019	0.036
	0.8	0.039	0.275	0.049	–0.001	0.268	–0.001	0.040
	1.0	0.064	0.288	0.064	0.018	0.280	0.018	0.046
	1.2	0.073	0.353	0.061	0.022	0.347	0.018	0.051
	1.4	0.057	0.360	0.041	0.002	0.358	0.001	0.056
Exchangeable								
	0.0	–0.005	0.283	*	–0.005	0.264	*	0.000
	0.2	–0.005	0.296	–0.025	–0.017	0.277	–0.089	0.013
	0.4	0.016	0.319	0.042	–0.008	0.304	–0.019	0.025
	0.6	0.003	0.334	0.005	–0.029	0.322	–0.048	0.033
	0.8	0.043	0.347	0.054	0.003	0.337	0.004	0.041
	1.0	0.034	0.353	0.034	–0.012	0.344	–0.012	0.046
	1.2	0.061	0.420	0.051	0.010	0.413	0.008	0.051
	1.4	0.053	0.422	0.038	–0.000	0.419	–0.000	0.054

Sometimes, the nature of the observations may be useful in determining the form of the true correlation structure, as in a family cohort study where the correlation of any two family members is often the same for the whole family and thus leads to an exchangeable structure. For a fixed sample size, use of true correlation in the GEE estimation enjoys several desirable statistical properties, such as small variance and high power. Unfortunately, the relation between power and bias is not monotonic at the micro level in the group sequential method. Consider the logistic regression model again and assume that the true exchangeable correlation is used in the GEE estimation. The information fractions for the uniform monitoring using the independence correlation in the GEE estimation are 0.25, 0.50, 0.75 and 1, and the corresponding O'Brien-Fleming boundaries are 4.332, 2.963, 2.359 and 2.014.

The information fractions using the true exchangeable structure are 0.596, 0.816, 0.930 and 1. Using the Brownian motion approximation, the O'Brien-Fleming boundaries are 2.679, 2.259, 2.143 and 2.091. Since observations from a single subject are all correlated, the two observations obtained at the first interim analysis contain information about observations not collected, and the amount of statistical information reaches almost 60% of the total available. Thus the O'Brien-Fleming boundary for the first monitoring is only 2.679.

Figure 5 shows the bias of the original and bias-corrected sequential estimates using the true correlation structure in both the GEE estimation and the Brownian motion approximation. Both types of bias are much larger than their counterpart in Figure 4. Table 2 shows power and average stopping time (AST) using the true exchangeable and the independence working correlation structures. Obviously, the independence structure has a larger average stopping time (AST) because of its larger boundaries. But the independence structure gives almost the same power as the true correlation structure. Thus using the independence structure only delays the stopping time of the trial. Therefore, bias using the independence working correlation is smaller than that using the true exchangeable correlation, as indicated in Figures 4 and 5.

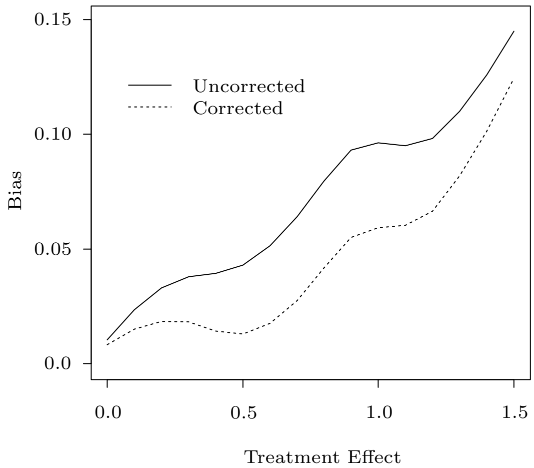

Figure 5. Bias reduction using O'Brien-Fleming's boundaries and the true exchangeable correlation structure for logistic regression with $\rho = 0.3$.

5. Discussion

This paper examines the bias of the group sequential estimate of treatment effect with correlated data using GEE estimation. Because the true correlation structure is unknown in most practical cases, our results focus on the use of the independence working correlation structure. Our simulation results are very close to what were reported for independent observations by others. With the independence working correlation model, the influence of the true correlation structure is determined by the overall level of correlation of observations from individual subjects. The higher the overall level of correlation, the smaller the bias of the sequential estimate. For the four correlation structures we examined, only the true exchangeable correlation structure distinguishes itself from other structures by having smaller bias.

Table 2. Power and average stopping time (AST) of sequential analysis with O'Brien-Fleming's boundaries calculated using the independence working correlation or the true exchangeable correlation structure for logistic regression with $\rho = 0.3$.

	γ	0.0	0.2	0.40	0.60	0.80	1.0	1.2	1.4
Independence	Power	0.042	0.097	0.269	0.466	0.709	0.838	0.926	0.971
	AST	3.977	3.951	3.805	3.578	3.217	2.897	2.582	2.372
Exchangeable	Power	0.053	0.093	0.248	0.480	0.690	0.865	0.933	0.977
	AST	3.925	3.845	3.580	3.179	2.677	2.126	1.836	1.506

Tables 1 and 2 present results using the O'Brien-Fleming's boundaries for logistic regressions. Similar simulations were also performed for Pocock's boundaries, and the results were similar. These are summarized in Qu and DeMets (1998).

We showed the magnitude of bias reduction for the logistic regression with equally spaced observations. However, with a fixed set of boundaries $\{c_k\}$ and for a given set of information fractions $\{t_i^*\}$ and ψ_K, the level of reduction is irrelevant to the model assumption since distribution (2) is determined completely by $\{t_i^*\}$ and ψ_K. Thus, the results we observed reflect the performance of this method in general. For most realistic situations, the O'Brien-Fleming type boundaries give relatively small bias compared to the Pocock type boundaries. Also, given the information fractions, the amount of reduction using (2) is positively related to the magnitude of the original bias in general.

Acknowledgement

Qu's work was supported by grant CA-09565 from the National Cancer Institute. We thank the Associate Editor and a referee for a very thorough review of the paper, and for their constructive suggestions.

References

DeMets, D. L. and Lan, K. K. G. (1994). Interim analysis: the alpha-spending function approach. *Statist. Medicine* **13**, 1341-1352.

Emerson, S. and Fleming, T. R. (1990). Parameter estimation following group sequential hypothesis testing. *Biometrika* **77**, 875-892.

Gange, S. J. and DeMets, D. L. (1996). Sequential monitoring of clinical trials with correlated responses. *Biometrika* **83**, 157-167.

Hughes, M. D. and Pocock, S. J. (1988). Stopping rules and estimation problems in clinical trials. *Statist. Medicine* **7**, 1231-1242.

Hughes, M. D., Freedman, L. and Pocock, S. J. (1992). The impact of stopping rules on heterogeneity of results in overview of clinical trials. *Biometrics* **48**, 41-53.

Laird, N. M. and Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics* **70**, 659-663.

Lan, K. K. G. and DeMets, D. L. (1983). Discrete sequential boundaries for clinical trials. *Biometrika* **70**, 659-663.

Lan, K. K. G. and Zucker, D. (1993). Sequential monitoring of clinical trials: the role of information in Brownian motion. *Statist. Medicine* **12**, 753-765.

Lan, K. K. G., Reboussin, D. M. and DeMets, D. L. (1994). Information and information fractions for design and sequential monitoring of clinical trials. *Comm. Statist. Part A-Theory and Methods* **23**, 403-420.

Lee, J. W. (1994). Group sequential testing in clinical trials with multivariate observations: a review. *Statist. Medicine* **13**, 101-111.

Lee, J. W. and DeMets, D. L. (1991). Sequential comparison of change with repeated measurement data. *J. Amer. Statist. Assoc.* **86**, 757-62.

Liang, K. Y. and Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika* **73**, 13-22.

Li, Z. Q. and DeMets, D. L. (1996). On the bias of Brownian motion drift following group sequential tests. Technical Report **103**, Department of Biostatistics, University of Wisconson-Madison.

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

Pinheiro, J. C. and DeMets, D. L. (1997). Estimating and reducing bias in group sequential designs with Gaussian independent increment structure. *Biometrika* **84**, 751-763.

Pocock, S. J. (1977). Group sequential methods in the design and analyses of clinical trials. *Biometrika* **64**, 191-199.

Qu, P. Q. and DeMets, D. L. (1998). On the bias correction in group sequential analysis with correlated data. Technical Report **130**, Department of Biostatistics, University of Wisconsin-Madison.

Reboussin, D. M., DeMets, D. L., Kim, K. and Lan, K. K. G. (1992). Programs for computing group sequential bounds using the Lan-Demets method. Technical Report **60**, Department of Biostatistics, University of Wisconsin-Madison.

Schervish, M. J. (1984). Algorithm as 195. Multivariate normal probability with error bound. *Appl. Statist.* **33**, 81-94.

Slud, E. and Wei, L. J. (1982). Two sample repeated significance tests based on the modified Wilcoxon statistic. *J. Amer. Statist. Assoc.* **77**, 862-868.

Wei, L. J., Su, J. Q. and Lachin, J. M. (1990). Interim analyses with repeated measurements in a sequential clinical trial. *Biometrika* **77**, 359-264.

Whitehead, J. (1986). On the bias of maximum likelihood estimation following a sequential test. *Biometrika* **73**, 573-581.

Wu, M. C. and Lan, K. K. Gordon (1992). Sequential monitoring for comparison of changes in a response variable in clinical studies. *Biometrics* **48**, 765-779.

Global Biometrics and Clinical Research, RWJPRI, 920 Route 202, P.O. Box 300, Raritan, New Jersey 08869, U.S.A.

E-mail: RQu@prius.jnj.com

Department of Biostatistics, University of Wisconsin-Madison, 600 Highland Avenue, Madison, Wisconsin 53792, U.S.A.

(Received September 1997; accepted June 1999)