King Fahd University of Petroleum & Minerals

DEPARTMENT OF MATHEMATICAL SCIENCES

Technical Report Series

TR 419

Jan 2011

Group Sequential Methods Based on Ranked Set Samples

H. A. Muttlak and E. Al-Sawi

Group Sequential Methods Based on Ranked Set Samples

A. Hussein
Department of Mathematics and Statistics, University of Windsor, Windsor, Ontario, Canada

H. A. Muttlak* and E. Al-Sawi

Department of Mathematics and Statistics, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia
*email: hamstat@kfupm.edu.sa

SUMMARY: Ranked set sampling schemes were originally proposed to increase efficiency in estimation. On the other hand, group sequential methods provide substantial savings in sample and enable us to make decisions as early as possible. In this manuscript, we intend to combine the benefits of the two methodologies. We propose group sequential tests for one and two population means under ranked set sampling. We compare the power, average sample sizes and type I errors of the proposed tests to those of group sequential tests based on simple random sampling schemes. We illustrate the utility of the method by using data from HIV trial.

KEY WORDS: Auxiliary information; Group sequential methods; Ranked set sampling; Two-sample hypotheses.

1. Introduction

Group sequential testing methods are often used in medical clinical trials in order to attain ethical, economical and administrative benefits. The most important benefit is the ethical one wherein, for instance, decisions regarding which of the two competitive treatments better can be made early and therefore, the patients can be duly removed from the inferior treatment arm. In the group sequential testing methodology, groups of observations are taken at each interim analysis (each time the data are examined for testing the hypotheses of interest) and, usually, a fixed sample statistic is computed at each interim analysis using the cumulative sample collected thus far. The first group sequential testing method was formalized by Pocock (1977) who dealt with the two-sample independent normal data and a composite alternative hypothesis about the mean difference. Pocock used equal group sizes and constant boundaries to satisfy pre-specified type I & II error probabilities. Later, O'Brien and

Fleming (1979) re-considered the same problem with square root boundaries and this method has the advantage of not rushing to reject the null hypothesis at early stages when the number of collected observations is small. The condition of equal group sizes was later removed by Lan and DeMets (1983) who introduced the so-called α-spending function to build the boundaries. Recently, sequential and group sequential methods have been extended in many directions and applied in various scientific areas of research; see for example Hussein and Carriere (2005) and Cui et al. (2009).

On the other hand, the sampling method coined as ranked set sampling (RSS) is an often more efficient method in many estimation and/or testing setups than the usual simple random sampling (SRS). In the simplest RSS procedure as suggested by McIntyre (1952), the investigator selects a SRS of $k \times k$ units from the population and randomly divides them into k subsets. Each subset is then ordered with respect to the quantity of interest without measuring it. This can be done by various ways such as by visual inspection or through ordering of concomitant variable. From the first set, the first order statistic is selected (i.e., the unit with the smallest measurement), whereas from the second subset the second order statistic is selected and so forth. Thus one will ultimately obtain a sample of size k and the entire procedure is repeated m times (called m cycles). Thus the final sample size has mk observations. In the context of hypothesis testing, Abu-Dayyeh and Muttlak (1996) and Muttlak and Abu-Dayyeh (1998) discussed testing under RSS for the mean of a normal and scale parameters of the exponential distribution. Takahasi and Wakimoto (1968) supplied the necessary mathematical theory. Dell and Clutter (1972) studied the case in which the ranking may not be perfect: i.e. there are errors in ranking the unit with respect to variable of interest. Recently, interest has been shown in the RSS by a number of investigators, for example Ozturk and Balakrishnan (2009) and Chen (2007).

The RSS method maybe used as an efficient way of incorporating auxiliary information at the design stage of a clinical trial. In many medical clinical trials, the variable of interest (X) has often concomitant (auxiliary) variables that are reasonably correlated with endpoint of interest and which can be used to rank patients prior to randomization. This is, in particular, helpful when measuring the variable of interest is difficult or costly, while visual inspection or measuring the auxiliary variable is cost-free or much cheaper. Also, many clinics receive patients in batches and it is actually feasible to divide potential patients into ranked subsets prior to randomization. Incorporating auxiliary information at the design stage of a clinical

trial has been attempted by Lagakos (1977). For instance, baseline CD4 counts may be useful auxiliary information in AIDS clinical trials. In clinical trials comparing two chemotherapies for the remission from leukemia, Karnofsky score or even demographic variables such as age or gender may serve as auxiliary variables.

This manuscript has, therefore, the objective of combining the benefits of group sequential methods and those of the RSS when testing hypothesis about the means of one and two populations. The use of RSS scheme would increase efficiency and hence, would result in procedures that stop much earlier than the usual group sequential procedures based on simple random samples. The rest of the manuscript is organized as follows. In Sections 2 and 3, we propose general RSS group sequential methods for one- and two-sample testing of hypothesis and we show that the score test statistics are asymptotically equivalent to discrete standard Brownian motion processes. In these sections, we illustrate how the efficiency of group sequential methods can, at times, be doubled by the use of RSS schemes. In Section 3, we give Monte Carlo simulations to compare the performance of the proposed methods and those of the usual SRS-based group sequential methods. In Section 4, we illustrate the utility of the new procedures by using data from a controlled clinical trial that compared two treatments for HIV infected persons (Hammer et. al. (1997)).

2. One-sample RSS Group Sequential Methods

2.1. The proposed method

In general, there are several types of ranked set sampling (RSS) schemes in the literature. However, in this manuscript we use the balanced RSS method, see McIntyre (1952). Suppose that we are interesting in group sequential testing of

$$H_0 : \mu = \mu_0 \text{ vs } H_1 : \mu \neq \mu_0 \tag{1}$$

based on a balanced RSS.

In order to conduct a RSS group sequential test of the hypotheses in (1), we need to decide the ranked set size, k , the overall type I error of the test, α , the analysis times, $0 \le t_1 \le t_2 \le ..., t_L \le 1$ and the α -spending function, $\alpha(t)$ (or, in other words, the boundary type to be used). The analysis times are usually on information scale, in the sense that they represent the fraction of information collected up to the l th interim analysis out of the total information to be collected at the end of the study. At

the 1st interim analysis, $l=1$, we select k sets, each consisting of k randomly selected subjects. The subjects in each set are then ranked with respect to the variable of interest without measuring the variable of interest. This can be done usually using any cost-free method, or using a concomitant variable. After ranking, data is collected only of the first subject (with smallest rank) of the first set, the 2nd smallest in the second set, and the largest from the last set. Thus we obtain measurements of the outcome $x_{[r]1}$ for $r=1,...,k$. This operation is then repeated m_1 times (cycles) until the sample size required to conduct the first interim analysis is obtained i.e. $n_1 = km_1$. At this point, a statistic Z_1 is computed, based on the sample n_1 collected thus far and the null hypothesis is rejected if $|Z_1| > c_1$, where c_1 is critical value to be computed by numerical integration as explained later. At the l th interim analysis, $l=2,...,L$, we repeat the above operation by randomly selecting k sets of k subjects and perform the RSS method. This operation is then repeated $(m_l - m_{l-1})$ times (cycles), for m_l such that the cumulative sample size at the l th interim analysis is $n_l = km_l$. Using the observations collected thus far, $x_{[r]j}$ for $r=1,...,k$ and $j=1,...,m_l$, we compute the test statistic Z_l using equation (2) below and reject the null hypothesis and stop testing if $|Z_l| > c_l$, where c_l is computed recursively from equation (7). At the last interim analysis, L, when the sample recruited is $N_L^{RSS} = km_L$, if still $|Z_L| > c_L$, then we stop testing and fail to reject H_0.

We notice here that we are not requiring equal group sizes. In other words, collecting the same sample size between each two consecutive interim analysis is not a requirement. Only the ranked set size, k, must be fixed over all analysis times, whereas the sample size at each interim analysis can vary, because the number of cycles, $(m_l - m_{l-1})$ is free.

In this manuscript, we use the sequence of test statistics $(Z_1, Z_2, ..., Z_L)$ given by

$$Z_l = \frac{\bar{x}_l^{rss} - \mu_0}{\sqrt{\hat{\tau}_l^2 / m_l}}, \tag{2}$$

where $\bar{x}_l^{rss} = \frac{1}{m_l k} \sum_{r=1}^{k} \sum_{j=1}^{m_l} x_{[r]j}$, $\hat{\tau}_l^2 = \sum_{r=1}^{k} \hat{\sigma}_{[r]l}^2$, and $\hat{\sigma}_{[r]l}^2$ is the estimated variance of the r th order statistic for a sample of size k see (Chen et al. (2004)),

$$\hat{\sigma}^2_{[r]l} = \frac{1}{m_l}\sum_{j=1}^{m_l}(x_{[r]j} - \bar{x}_{[r]l})^2, \tag{3}$$

$$\bar{x}_{[r]l} = \frac{1}{m_l}\sum_{j=1}^{m_l} x_{[r]j}.$$

Notice that these estimators are all based on the cumulative data up to the l th interim analysis.

The following result allows us to use the Brownian motion approximation so that the usual group sequential boundaries of Pocock (1977), O'Brien and Fleming (1979) and boundaries based on the α-spending function approach of Lan and DeMets (1983) can be adopted.

Theorem 1. Suppose the variance of the population, σ^2, is finite, and the ranking mechanism is consistent (Chen et al. (2004)), i.e.,

$$\mu = \frac{1}{k}\sum_{r=1}^{k}\mu_{[r]} \tag{4}$$

then,

$$(Z_1,\ldots,Z_L) \xrightarrow{D} (\frac{B(t_1)}{\sqrt{t_1}},\ldots,\frac{B(t_L)}{\sqrt{t_L}},) \text{ as } m_L \to \infty \text{ and } \frac{m_l}{m_L} \to \lambda_l < \infty,$$

where $B(t_l)$ is a standard Brownian motion computed at time t_l.

Proof. *Since* $\hat{\tau}^2_l$ is a consistent estimator of τ^2 as $m_l \to \infty$, by Slutsky's theorem (Lehmann (1999)), all we need is to show that this result is true for the vector $\mathbf{S}^{m_L} = (S_1^{m_L}, S_2^{m_L}, ..., S_L^{m_L})$, where

$$S_l^{m_L} = \frac{\bar{x}_l^{rss} - \mu_0}{\sqrt{\tau^2 / m_l}}. \tag{5}$$

This, in turn, amounts to showing that $\mathbf{S}^{m_L}$ has asymptotically multivariate normal with mean zero and covariance matrix with the $(i, j)th$ element given by $\sqrt{t_j / t_i}$ for $j \le i$. To this end, first we notice that, because of (4), S^{m_L} can be re-written as

$$\mathbf{S}^{m_L} = \mathbf{C}^{m_L}\mathbf{V}^{m_L},$$

where $\mathbf{V}^{m_L} = (V_1^{m_L}, V_2^{m_L}, ..., V_L^{m_L})$ is a random vector whose components are mutually independent and each is a standardized sum of i.i.d random variables,

$$V_l^{m_L} = \frac{1}{\sqrt{(m_l - m_{l-1})\tau^2}} \sum_{j=m_{l-1}+1}^{m_l} y_j,$$

and

$$y_j = \frac{1}{k}\sum_{r=1}^{k}(x_{[r]j} - \mu_{[r]}^0) \tag{6}$$

are independent random variables with $E[Y_j] = 0$ and

$$Var(Y_j) = \frac{1}{k^2}\sum_{r=1}^{k}\sigma_{[r]}^2$$

with $\sigma_{[r]}^2 = Var(X_{[r]})$ and $\mu_{[r]}^0 = E[X_{[r]} \mid H_0]$. It can be shown that

$$\tau^2 = \frac{1}{k}\sigma^2 - \frac{1}{k^2}\sum_{r=1}^{k}(\mu_{[r]}^0 - \mu_0)^2 \le \frac{\sigma^2}{k}$$

where $\sigma^2 = Var(X)$. Therefore, by the assumption of finiteness of σ^2, it also follows that, $\tau^2 < \infty$ for any fixed k. Thus, as m_l and $m_L \to \infty$, by the standard central limit theorems, we have that $V_l^{m_L} \xrightarrow{D} Z_l$, a standard normal random variable for all $l = 1, 2, ..., L$. Hence, $\mathbf{V}^{m_L} \xrightarrow{D} \mathbf{Z}$, where $\mathbf{Z}$ has L-variate standard normal distribution with identity covariance matrix. On the other hand, $\mathbf{C}^{m_L}$ is a $L \times L$ matrix with elements

$$C_{i,j}^{m_L} = \left(\frac{(m_j - m_{j-1})\tau^2}{m_i\tau^2}\right)^{1/2} = \left(\frac{(I_j/I_L - I_{j-1}/I_L)}{I_i/I_L}\right)^{1/2} \to \left(\frac{(t_j - t_{j-1})}{t_i}\right)^{1/2} = C_{i,j}^{L}$$

for $i \ge j$ and zero otherwise. Here we are assuming that analysis times are limits of the information fractions, i.e., $\frac{m_j/\tau^2}{m_L/\tau^2} = I_j/I_L \to t_j$. Thus, by multivariate Slutsky's theorem, we can assert that $\mathbf{C}^{m_L}\mathbf{V}^{m_L} \xrightarrow{D} \mathbf{C}^{L}\mathbf{Z}$ with $Cov(\mathbf{C}^L\mathbf{Z}) = (\mathbf{C}^L)(\mathbf{C}^L)^t = \left(\sqrt{t_j/t_i}\right)_{i,j}$ and hence, the desired result follows.

The group sequential boundaries of the study can now be computed by using common software or tables available from Jennison and Turnbull (2000). In particular, the boundaries can be computed recursively through an α-spending function, $\alpha(t)$. In general, the function $\alpha(t)$ determines the cumulative portion of the overall type I error that has been spent on or prior to the l th interim analysis and, therefore, it is a nondecreasing function such that $\alpha(0) = 0$ and $\alpha(1) = \alpha$ (see Lan and DeMets (1983)). The boundaries, $c_1, c_2, \ldots, c_L$, for monitoring the Z_l statistic at the

interim analysis times $t_1,\ldots,t_L$ and for the hypotheses in (1) can be computed from the equations

$$\alpha(t_1) = P\{|Z_1| \geq c_1 \mid H_0\} \approx P\left\{\left|\frac{B(t_1)}{\sqrt{t_1}}\right| \geq c_1\right\} = 2(1-\Phi(c_1))$$

$$\alpha(t_l) = \sum_{i=1}^{l} P\{|Z_1| < c_1,\ldots,|Z_{i-1}| < c_{i-1}, |Z_i| \geq c_i \mid H_0\}$$

$$\approx \sum_{i=1}^{l} P\left\{\left|\frac{B(t_1)}{\sqrt{t_1}}\right| < c_1,\ldots,\left|\frac{B(t_{i-1})}{\sqrt{t_{i-1}}}\right| < c_{i-1}, \left|\frac{B(t_i)}{\sqrt{t_i}}\right| \geq c_i\right\}, \quad (7)$$

where $l = 2,3,\ldots,L$.

2.2. Efficiency of the RSS group sequential in the normal case

In planning for a group sequential study, we need to determine the sample size needed to attain certain pre-specified power. To this end, we assume that the cumulative information up to the l th interim analysis for the RSS group sequential procedure is $I_l = m_l / \tau_l^2$, i.e., the inverse of the variance of the estimator $\bar{X}_l^{rss}$. As before, the information times are assumed to be $t_l \approx I_l / I_L$. Thus, in designing a RSS group sequential study, we compute the drift of the Brownian motion, $d = \sqrt{I_L}\,\delta$, where $\delta = |\mu_0 - \mu_a|$, under a presumed alternative hypothesis $H_a : \mu = \mu_a$. Such a drift can be obtained by an obvious modification of the integrals in (7) based only on the power specification $(1-\beta)$ and without the knowledge of the alternative value or even the maximum information at the end of the study.

Once the drift has been computed, the required sample size is obtained by specifying the relationship between information, variance, sample size, and drift. In the RSS case, such a relationship is given by

$$I_L^{-1} = \frac{\tau^2}{m_L} = \frac{1}{m_L}\left[\frac{1}{k}\sigma^2 - \frac{1}{k^2}\sum_{r=1}^{k}(\mu_{[r]}^0 - \mu_0)^2\right] \quad (8)$$

where $\mu_{[r]}$ and $\mu_{[r]}^0$, respectively, are the mean of the r th order statistic based on sample of size k and its value under the null hypothesis in (1). Computing the expected value of the r th order statistic is, therefore, required for obtaining sample size through the drift parameter. On the other hand, if we assume normality, by using the well known relationship between moments of order statistics of the normal and

their standard normal counterparts, the above equation simplifies to

$$I_L^{-1} = \frac{\tau^2}{m_L} = \frac{\sigma^2}{km_L}\gamma_k, \tag{9}$$

where

$$\gamma_k = 1 - \frac{1}{k}\sum_{r=1}^{k} \mu_{z[r]}^2$$

and $\mu_{z[r]}$ is the mean of the r th standard normal order statistic based on a sample of size k ,

$$\int_{-\infty}^{\infty} \frac{k!}{(r-1)!(k-r)!} z[\Phi(z)]^{r-1}[1-\Phi(z)]^{k-r}\varphi(z)dz. \tag{10}$$

The values of this integral are tabulated by many authors, see for example Harter and Balakrishnan (1996). The factor γ_k is such that $0 \le \gamma_k \le 1$, $\gamma_1 = 1$, and $\gamma_k \to 0$ as $k \to \infty$. The first few values of the integral are reported in Table1.

Table 1

γ_k for different values of k

k	1	2	3	4	5	6	7	8	9	10
γ_k	1.0000	0.6817	0.5225	0.4261	0.3610	0.3139	0.2782	0.2501	0.2273	0.2086

Thus, the relationship between sample size required by the RSS group sequential design and that required by the usual SRS group sequential designs is

$$N_L^{RSS} = \left(\frac{d^2\sigma^2}{\delta^2}\right)\gamma_k = N_L^{SRS}\gamma_k \tag{11}$$

where N_L^{SRS} is the maximum sample size required by a SRS group sequential for a fixed power $(1-\beta)$ and hence, a fixed drift δ .

Therefore, the maximal sample size required by an RSS group sequential is γ_k times smaller than that required by an equally powerful group sequential procedure based on the usual SRS scheme. If we use RSS group sequential procedure with $k = 3$, we obtain almost 52.25% reductions in the sample required by equally powerful group sequential procedures based on SRS. It is also clear that a RSS group sequential procedure with $k = 3$ for testing H_0 in (1) can be designed by simply using the design of an equivalent, in power, SRS group sequential design and then

recruiting only half of the maximal sample required by such a design.

The extra subjects that we need to assess for ranking purposes are compensated by the saving attained through the RSS design in terms of the actual number of subjects recruited and measured. To be more specific, if a SRS group sequential with a certain, pre-fixed, power, type I error, and L interim analysis, requires a sample of N_L^{SRS} subjects to be recruited and measured, then for an equivalent RSS group sequential procedure all we need to recruit and measure is $N_L^{RSS} = \gamma_k N_L^{SRS}$ fraction of the subjects; However, ranking of an extra $N_T^{RSS} = k\,\gamma_k N_L^{SRS}$ is required (by means of a concomitant variable, such as a biomarker, age, or a cheap variable such as blood pressure). For example, if SRS group sequential requires $N_L^{SRS} = 200$, then for an RSS with $k = 2$, we need to rank $N_T{}^{RSS} = 273$ subjects but recruit and measure the outcome only for $N_L{}^{RSS} = 136$ subjects, whereas for $k = 3$, these numbers are 314 and 105, respectively. This is an enormous saving in terms of maximal sample as well as average sample required for detecting the alternative hypotheses.

3. Two-sample Extensions

In the case of two-populations, one would independently obtain two RSS samples, denoted by $x_{1[r]j}$ and $x_{2[r]j}$, respectively.

The extension of the results in earlier sections, to the case of two-sample problems, is quite straight forward. Here we propose an RSS scheme whereby an independent balanced RSS sample is drawn from each population as described above in the one-sample case. Thus, let m_{il} be the cumulative number of cycles by the lth interim analysis from the two populations being compared, where $i = 1, 2$ and $l = 1, 2, ..., L$ as before. The sequence of RSS-based test statistics for testing

$$H_0 : \mu_1 - \mu_2 = 0;\ H_a : \mu_1 - \mu_2 \neq 0$$

can be written as

$$Z_l = \frac{\bar{X}_{1l}^{rss} - \bar{X}_{2l}^{rss}}{\sqrt{\hat{\tau}_{1l}^2 / m_{1l} + \hat{\tau}_{2l}^2 / m_{2l}}}, \tag{12}$$

where the various quantities are defined as in the one-sample case. Again, we can prove that

$$\mathbf{Z}^{m_L} = (Z_1, ..., Z_L) \xrightarrow{D} (\frac{B(t_1)}{\sqrt{t_1}}, \ldots, \frac{B(t_L)}{\sqrt{t_L}}),$$

as $m_{1L}, m_{2L} \to \infty$ and $m_{1L} \Big/ m_{2L} \to \lambda_L < \infty$, $m_{1l} \Big/ m_{1L} \to \lambda_{1l} < \infty$, $m_{2l} \Big/ m_{2L} \to \lambda_{2l} < \infty$.

This can be accomplished again, by proving the result for the vector $\mathbf{S}^{m_L} = (S_1^{m_L}, S_2^{m_L}, ..., S_L^{m_L}) = \mathbf{Z}^{m_L} + o_p(1)$. Such vector can be represented as

$$\mathbf{S}^{m_L} = \mathbf{C}_1^{m_L} \mathbf{V}_1^{m_L} - \mathbf{C}_2^{m_L} \mathbf{V}_2^{m_L},$$

where $\mathbf{V}_i^{m_L}$ for $i = 1, 2$ is a vector of L mutually independent random variables of the form

$$V_{il}^{m_L} = \left(\frac{1}{(m_{il} - m_{i\,(l-1)})\tau_i^2} \right)^{1/2} \sum_{j=m_{i\,(l-1)}}^{m_{il}} y_{ij},$$

$\mathbf{C}_i^{m_L}$ are $L \times L$ matrices with elements

$$C_{i(l,l')}^{m_L} = \frac{\left((m_{il'} - m_{i(l'-1)})\tau_i^2 \right)^{1/2}}{m_{il} \left(\tau_1^2 / m_{1l} + \tau_2^2 / m_{2l} \right)^{1/2}}$$

for $l' \le l$ and zero otherwise, and y_{ij} as defined in (6). Assume that the design is information-balanced in the sense that the information accrued in the two treatment arms at each interim analysis are equal. That is,

$$I_{il} = \frac{m_{il}}{\tau_i^2} = I_l \text{ for } i = 1, 2.$$

Therefore, we can easily see that

$$C_{i(l,l')}^{m_L} = \frac{((m_{il'} - m_{i(l'-1)})\tau_i^2)^{1/2}}{m_{il} \left(\tau_1^2 m_{1l} + \tau_2^2 m_{2l} \right)^{1/2}} = \frac{1}{\sqrt{2}} \left(\frac{I_{il'} / I_L - I_{i(l'-1)} / I_l}{I_l / I_L} \right)^{1/2} \to \frac{1}{\sqrt{2}} \left(\frac{t_{l'} - t_{(l'-1)}}{t_l} \right)^{1/2} = \frac{1}{\sqrt{2}} C_{(l,l')}^L \quad (13)$$

as $m_{iL} \to \infty$. The vectors, $\mathbf{V}_{il}^{m_L}$ converge in distribution to independent multivariate standard normal variables, $\mathbf{Z}_1$ and $\mathbf{Z}_2$, as in the one-sample case and hence, by multivariate Slutsky's theorem, we have

$$\mathbf{S}^{m_L} = \mathbf{C}_1^{m_L} \mathbf{V}_1^{m_L} - \mathbf{C}_2^{m_L} \mathbf{V}_2^{m_L} \xrightarrow{D} \frac{1}{\sqrt{2}} \mathbf{C}^L \mathbf{Z}_1 - \frac{1}{\sqrt{2}} \mathbf{C}^L \mathbf{Z}_2.$$

Therefore, the asymptotic distribution of $\mathbf{S}^{m_L}$ is multivariate standard normal with

covariance matrix $\boldsymbol{\Sigma} = \frac{1}{2}\mathbf{C}^L(\mathbf{C}^L)^T + \frac{1}{2}\mathbf{C}^L(\mathbf{C}^L)^T = \mathbf{C}^L(\mathbf{C}^L)^T$ whose elements are $\sigma_{l,l'} = \sqrt{t_{l'}/t_l}$ for $l' \le l$.

4. Numerical Studies

4.1. Simulation Studies

Here we consider fixed total sample sizes of $N = 60, 120, 180, 380, 600$, set sizes $k = 2, 3$, number of interim analysis $L = 1, 2, 4, 5, 10$, type I error $\alpha = 0.05$ with Pocock and O'Brien-Fleming boundaries. The hypotheses in (1) were considered with $\mu = 0, 0.1, 0.15, 0.2, 0.25$ and normal distributions with $\sigma^2 = 1$. In each scenario, we have used 10^4 Monte Carlo replications. The ranking was based on concomitant variable y which is also normally distributed and whose correlation with the variable of interest x was varied over $\rho = 0, 0.25, 0.50, 0.75, 1$. To conserve space, we only reported results for O'Brien-Fleming boundaries and $\mu = 0, 0.1, 0.2$ (Tables 2 - 7). From Tables 2 & 3, we can see that at large total sample sizes, both SRS and RSS methods maintain their nominal α. On the other hand, for small samples, such as $N = 60$, and small number of interim analysis, the RSS group sequential methods maintain their type I errors while their SRS counterparts have inflated type errors. When the number of analysis is large ($L = 1$) relative to the total sample size, then the RSS methods are more inflated than their SRS counterparts. The type I errors are not much affected by changing correlation with the concomitant variable. This argument carries on for both ranked set sizes $k = 2, 3$, although the case $k = 3$ is slightly better in maintaining type I errors.

From Tables 4-7, we see that the RSS group sequential methods have substantially higher powers than their SRS counterparts and require, on average, less sample size to detect the same alternatives. The power gain by using RSS can reach up to 28% (see Table 5 with $k = 3$, $L = 5$ and $N = 180$, $\rho = 1$). At low correlations such as $\rho = 0.25$, the gain in power and average sample size is not substantial. When correlation is zero (not reported here), the two methods are the same (as expected).

The simulations are also in support of the arguments in Section 2, whereby we stated that the RSS group sequential methods would require only γ_k fraction of the

total sample size required by an equally powerful SRS group sequential procedure. This can be seen by examining powers of the RSS with $N = 180$ and SRS with $N = 360$. These samples are approximately differing by $\gamma_3 = 0.5225$ and we can see that the powers of the RSS and SRS are quite close at these sample sizes, which confirms equation (11).

4.2. Clinical trial example

In this section, we illustrate the proposed methods by using a data set from the well known ACGT 320 controlled clinical trial (see Hammer et. al (1997)). This trial compared two treatments (namely, three-drug combination of IDV+ZDV+3TC (treatment 1) vs two-drug combination of ZDV+3TC (treatment 2)) in persons with HIV. The primary outcome was survival end-point (until development of AIDS or death). However, an important secondary end-point was CD4 count of the patients at weeks: 0(baseline), 4, 8, 24 and 40. Here found that for the CD4 end-point, the correlations between counts at baseline and week 4 were quite strong ($\rho_1 = 0.80$ and $\rho_1 = 0.86$, respectively, under treatments 1 & 2). Thus we used the CD4 counts at week 0 as our concomitant variable. A SRS-based O'Brien-Fleming design with L=4, significance level of 0.05 and power of 85% would require 420 patients on each arm in order to detect a minimum difference equal to the observed absolute difference of, approximately, 19 cells/mm^3 between the CD4 counts of the two treatment groups at week 4 as reported in Hammer et. al. (1997). By using simulations we , it turns out that an equally powerful O'Brien-Fleming RSS design with k=3 would require an effective total sample size of 300 whereas if k=2 it would require 340.

Using such designs, the available data set of about 532 patients on each arm, and the Z-statistic in equation (12), we monitored the above trial and found that the trial stops at interim analysis 2 under all designs. The different designs, their Z-statistics and corresponding monitoring boundaries as well as the sample size on each arm at the time the trials stopped are reported in Table 8. In summary, the RSS designs would enjoy the same power, but save as much as half of the required sample size.

5. Conclusions

In this manuscript, we proposed group sequential testing procedures based ranked set sampling for one- and two-population mean comparisons. We have shown analytically and by simulations that the proposed methods are more efficient than the usual group sequential methods based on simple random sample. We illustrated how the proposed methods could be applied in real-life by using data from HIV clinical trial example.

ACKNOWLEDGEMENTS

This work was supported by King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia under the Fast Track project # FT/ 100019.

REFERENCES

Abu-Dayyeh, W. and Muttlak, H. A. (1996). Using ranked set sampling for hypothesis tests on the scale parameter of the exponential and uniform distributions. *Pakistan Journal of Statist*ics, **12**,131-138.

Chen, Z., Zhidong B. and Sinha, Bimal K (2004). *Ranked Set Sampling*. Vol. 176 of Lecture Notes in Statistics. New York: Springer-Verlag. Theory and applications.

Cui, Y., Yuejiao, F. and Hussein, A. (2009). Group sequential testing of homogeneity in genetic linkage analysis. *Computational Statistical & Data Analysis*, **53**, 3630-3639.

Dell, T. R. and Clutter, J. L. (1972). Ranked set sampling theory with order statistics background. *Biometrics*, **28**, 545-553.

Emerson, S. S. and Banks P. L. (1994). Interpretation of a leukemia trial stopped early. In *Case Studies in Biometry* (Lange, N., Ryan, L ., Billard, L., Brillinger, D., Conquest, L. and Greenhouse, J. Eds.). New York: Wiley-Interscience.

Hammer, S. M., Squires, K. E., Hughes, M. D., Grimes, J. M., Demeter, L. M., Currier, J. S., Eron, J. J Jr., Feinberg, J. E., Balfour, H. H Jr., Deyton, L. R., Chodakewitz, J. A., Fisch, M. A. (1997). A controlled trial of two nucleosides analogues plus indinavir in persons with human immunodeficiency virus infection and CD4 cell counts of 200 per cubic millimiter or less. New Eng. J. Med. **337,** 725–733.

Harter, H. L. and Balakrishnan, N. (1996). *CRC Handbook of tables for the use of order*

statistics in estimation, Boca Raton: CRC press.

Hussein, A and Carriere, K. C. (2005). On group sequential procedures under variance heterogeneity. *Statistical Methods in Medical Research*, **14**, 121-128.

Jennison, C. and Turnbull, B. W. (2000). *Group Sequential Methods with Application to Clinical Trials*. Florida: Chapman & Hall/CRC.

Lagakos, S. W. (1977). Using auxiliary variables for improved estimates of survival time. *Biometrics*, **33**, 399-404.

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

Lehmann, E. (1999). *Elements of Large Sample Theory*. New York: Springer.

McIntyre, G. A. (1952). A method for unbiased selective sampling, using, ranked sets. *Australian Journal of Agricultural Research*, **3**, 385-390

Muttlak, H. A. and Abu-Dayyeh, W. A. (1997). Testing hypotheses for the normal distribution using ranked set sampling. *Journal of Information & Optimization Sciences*, **19**, 1-11.

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

Ozturk, O. and Balakrishnan, N. (2009). An exact control-versus-treatment comparison test based on ranked set samples. *Biometrics*, **65**, 1213-1222.

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

Takahasi, K. and Wakimoto, K. (1968). On the unbiased estimates of the population mean based on the sample stratified by means of ordering. *Annals of Institute of Statistical Mathematics*, **20**, 1-31.

Table 2. Simulated type I error for RSS, *Prss* , and percentage difference in type I errors between RSS and SRS, *Pd* , for total samples N , interim analysis, L , set size $k = 2$, correlation with concomitant variable, ρ , under O'Brien-Fleming boundaries.

		ρ									
		0.00		0.25		0.5		0.75		1	
N	L	Prss	Pd	Prss	Pd	Prss	Pd	Prss	Pd	Prss	Pd
60	1	0.0534	-0.3%	0.0551	-0.1%	0.0541	-0.2%	0.0544	-0.2%	0.0544	-0.2%
	2	0.0593	0.3%	0.0590	0.3%	0.0559	-0.1%	0.0571	0.1%	0.0568	0.0%
	4	0.0600	-0.1%	0.0588	-0.2%	0.0617	0.1%	0.0606	0.0%	0.0604	0.0%
	5	0.0614	0.1%	0.0596	-0.1%	0.0600	-0.1%	0.0595	-0.1%	0.0596	-0.1%
	10	0.0625	0.1%	0.0633	0.2%	0.0635	0.2%	0.0644	0.3%	0.0654	0.4%
120	1	0.0529	-0.1%	0.0528	-0.1%	0.0529	-0.1%	0.0534	-0.1%	0.0519	-0.2%
	2	0.0544	0.1%	0.0523	-0.1%	0.0519	-0.1%	0.0513	-0.2%	0.0510	-0.2%
	4	0.0558	0.2%	0.0550	0.2%	0.0541	0.1%	0.0546	0.1%	0.0549	0.2%
	5	0.0529	-0.3%	0.0546	-0.2%	0.0562	0.0%	0.0528	-0.3%	0.0534	-0.3%
	10	0.0551	0.0%	0.0548	0.0%	0.0543	-0.1%	0.0568	0.2%	0.0551	0.0%
180	1	0.0508	-0.1%	0.0511	0.0%	0.0509	0.0%	0.0511	0.0%	0.0503	-0.1%
	2	0.0501	-0.4%	0.0538	0.0%	0.0512	-0.3%	0.0525	-0.2%	0.0549	0.1%
	4	0.0537	0.1%	0.0530	0.1%	0.0539	0.2%	0.0524	0.0%	0.0524	0.0%
	5	0.0545	0.5%	0.0531	0.3%	0.0523	0.2%	0.0545	0.5%	0.0526	0.3%
	10	0.0518	-0.1%	0.0546	0.2%	0.0535	0.1%	0.0531	0.0%	0.0540	0.1%
360	1	0.0518	0.0%	0.0510	-0.1%	0.0526	0.0%	0.0507	-0.1%	0.0509	-0.1%
	2	0.0524	0.2%	0.0511	0.1%	0.0517	0.1%	0.0506	0.0%	0.0515	0.1%
	4	0.0520	0.1%	0.0524	0.1%	0.0507	0.0%	0.0528	0.2%	0.0507	0.0%
	5	0.0501	0.0%	0.0515	0.2%	0.0506	0.1%	0.0520	0.2%	0.0510	0.1%
	10	0.0512	0.0%	0.0524	0.1%	0.0516	0.0%	0.0507	-0.1%	0.0518	0.1%
600	1	0.0504	0.2%	0.0515	0.3%	0.0503	0.1%	0.0515	0.3%	0.0505	0.2%
	2	0.0506	0.0%	0.0516	0.1%	0.0490	-0.2%	0.0519	0.1%	0.0503	-0.1%
	4	0.0493	-0.1%	0.0519	0.1%	0.0525	0.2%	0.0514	0.1%	0.0522	0.2%
	5	0.0502	0.1%	0.0524	0.3%	0.0503	0.1%	0.0499	0.0%	0.0511	0.1%
	10	0.0503	0.3%	0.0507	0.4%	0.0513	0.4%	0.0519	0.5%	0.0515	0.5%

Table 3. Simulated type I error for RSS, *Prss* , and percentage difference in type I errors between RSS and SRS, *Pd* , for total samples N , interim analysis, L , set size $k = 3$, correlation with concomitant variable, ρ , under O'Brien-Fleming boundaries.

		ρ									
		0.00		0.25		0.5		0.75		1	
N	L	Prss	Pd	Prss	Pd	Prss	Pd	Prss	Pd	Prss	Pd
60	1	0.0553	-0.1%	0.0532	-0.3%	0.0549	-0.1%	0.0558	0.0%	0.0566	0.0%
	2	0.0584	0.2%	0.0584	0.2%	0.0579	0.1%	0.0576	0.1%	0.0571	0.1%
	4	0.0604	0.0%	0.0612	0.0%	0.0611	0.0%	0.0587	-0.2%	0.0598	-0.1%
	5	0.0620	0.1%	0.0594	-0.1%	0.0601	-0.1%	0.0614	0.1%	0.0596	-0.1%
	10	0.0699	0.9%	0.0696	0.8%	0.0698	0.9%	0.0698	0.9%	0.0689	0.8%
120	1	0.0524	-0.2%	0.0524	-0.2%	0.0519	-0.2%	0.0526	-0.1%	0.0519	-0.2%
	2	0.0533	0.0%	0.0545	0.1%	0.0554	0.2%	0.0522	-0.1%	0.0555	0.2%
	4	0.0551	0.2%	0.0553	0.2%	0.0554	0.2%	0.0542	0.1%	0.0540	0.1%
	5	0.0537	-0.3%	0.0544	-0.2%	0.0539	-0.2%	0.0555	-0.1%	0.0540	-0.2%
	10	0.0548	0.0%	0.0551	0.0%	0.0556	0.0%	0.0557	0.1%	0.0546	-0.1%
180	1	0.0508	-0.1%	0.0522	0.1%	0.0501	-0.1%	0.0524	0.1%	0.0531	0.2%
	2	0.0510	-0.3%	0.0533	-0.1%	0.0518	-0.2%	0.0518	-0.2%	0.0519	-0.2%
	4	0.0524	0.0%	0.0530	0.1%	0.0527	0.0%	0.0513	-0.1%	0.0534	0.1%
	5	0.0523	0.2%	0.0535	0.4%	0.0517	0.2%	0.0535	0.3%	0.0531	0.3%
	10	0.0537	0.1%	0.0525	0.0%	0.0527	0.0%	0.0526	0.0%	0.0543	0.2%
360	1	0.0505	-0.2%	0.0507	-0.1%	0.0506	-0.2%	0.0503	-0.2%	0.0507	-0.2%
	2	0.0515	0.1%	0.0526	0.2%	0.0500	-0.1%	0.0511	0.1%	0.0519	0.1%
	4	0.0524	0.1%	0.0516	0.0%	0.0538	0.3%	0.0514	0.0%	0.0514	0.0%
	5	0.0526	0.3%	0.0513	0.1%	0.0504	0.1%	0.0491	-0.1%	0.0518	0.2%
	10	0.0529	0.2%	0.0530	0.2%	0.0514	0.0%	0.0513	0.0%	0.0500	-0.1%
600	1	0.0514	0.3%	0.0502	0.1%	0.0500	0.1%	0.0501	0.1%	0.0497	0.1%
	2	0.0502	-0.1%	0.0495	-0.2%	0.0507	0.0%	0.0505	-0.1%	0.0502	-0.1%
	4	0.0530	0.3%	0.0517	0.1%	0.0501	0.0%	0.0503	0.0%	0.0522	0.2%
	5	0.0511	0.1%	0.0506	0.1%	0.0520	0.2%	0.0493	0.0%	0.0505	0.1%
	10	0.0495	0.3%	0.0497	0.3%	0.0512	0.4%	0.0503	0.4%	0.0498	0.3%

Table 4. Simulated power of group sequential RSS method, *Prss* , percentage power difference between RSS and SRS P_d , difference of average sample sizes between RSS and SRS as percentage of the total sample, eff, total sample size N , correlation with the concomitant variable ρ , number of interim analysis L , set size, $k = 2$, and alternative $H_a : \mu = 0.1$ under O'Brien-Fleming boundaries.

		ρ											
		0.25			0.5			0.75			1		
N	L	Prss	Pd	eff	Prss	Pd	eff	Prss	Pd	eff	Prss	Pd	eff
60	1	0.130	0%	0%	0.134	1%	0%	0.145	2%	0%	0.161	4%	0%
	2	0.130	0%	0%	0.133	1%	0%	0.147	2%	0%	0.163	4%	0%
	4	0.130	0%	0%	0.134	0%	0%	0.150	2%	1%	0.164	4%	1%
	5	0.134	0%	0%	0.137	1%	0%	0.143	1%	0%	0.164	3%	1%
	10	0.135	0%	0%	0.141	0%	0%	0.151	1%	0%	0.170	3%	1%
120	1	0.201	0%	0%	0.210	1%	0%	0.228	3%	0%	0.266	7%	0%
	2	0.201	0%	0%	0.207	1%	0%	0.231	3%	0%	0.268	7%	1%
	4	0.200	0%	0%	0.208	1%	0%	0.229	3%	1%	0.268	7%	1%
	5	0.199	0%	0%	0.210	1%	0%	0.224	3%	1%	0.264	7%	1%
	10	0.199	1%	0%	0.210	2%	0%	0.227	3%	1%	0.262	7%	2%
180	1	0.276	1%	0%	0.288	2%	0%	0.321	5%	0%	0.366	10%	0%
	2	0.274	0%	0%	0.285	1%	0%	0.318	4%	0%	0.371	10%	1%
	4	0.268	0%	0%	0.284	2%	0%	0.311	4%	1%	0.365	10%	2%
	5	0.270	1%	0%	0.284	2%	0%	0.312	5%	1%	0.363	10%	2%
	10	0.268	0%	0%	0.279	2%	0%	0.310	5%	1%	0.363	10%	3%
360	1	0.478	1%	0%	0.513	4%	0%	0.550	8%	0%	0.634	16%	0%
	2	0.481	1%	0%	0.500	3%	0%	0.551	8%	1%	0.631	16%	3%
	4	0.474	1%	0%	0.491	2%	1%	0.545	8%	2%	0.625	16%	4%
	5	0.473	1%	0%	0.495	3%	1%	0.546	8%	2%	0.622	16%	5%
	10	0.464	0%	0%	0.494	3%	1%	0.538	8%	2%	0.617	16%	5%
600	1	0.696	1%	0%	0.721	3%	0%	0.775	9%	0%	0.843	15%	0%
	2	0.697	1%	0%	0.721	4%	1%	0.768	9%	2%	0.839	16%	5%
	4	0.686	1%	0%	0.716	4%	1%	0.762	8%	3%	0.835	16%	7%
	5	0.682	0%	0%	0.711	3%	1%	0.760	8%	3%	0.834	16%	7%
	10	0.680	1%	0%	0.707	4%	2%	0.756	9%	4%	0.827	16%	8%

Table 5. Simulated power of group sequential RSS method, *Prss* , percentage power difference between RSS and SRS P_d , difference of average sample sizes between RSS and SRS as percentage of the total sample, eff, total sample size N , correlation with the concomitant variable ρ , number of interim analysis L , set size, $k = 3$, and alternative $H_a : \mu = 0.1$ under O'Brien-Fleming boundaries.

		ρ											
		0.25			0.5			0.75			1		
N	L	Prss	Pd	eff	Prss	Pd	eff	Prss	Pd	eff	Prss	Pd	eff
60	1	0.129	0%	0%	0.137	1%	0%	0.156	3%	0%	0.195	7%	0%
	2	0.132	1%	0%	0.138	1%	0%	0.156	3%	0%	0.195	7%	1%
	4	0.134	0%	0%	0.139	1%	0%	0.158	3%	1%	0.196	7%	1%
	5	0.134	0%	0%	0.139	1%	0%	0.158	3%	1%	0.200	7%	1%
	10	0.141	0%	1%	0.151	1%	1%	0.166	3%	1%	0.200	6%	2%
120	1	0.202	0%	0%	0.217	2%	0%	0.252	5%	0%	0.330	13%	0%
	2	0.202	1%	0%	0.218	2%	0%	0.252	6%	1%	0.332	14%	1%
	4	0.204	1%	0%	0.220	2%	0%	0.249	5%	1%	0.325	13%	2%
	5	0.203	1%	0%	0.216	2%	0%	0.247	5%	1%	0.327	13%	3%
	10	0.199	1%	0%	0.214	2%	0%	0.247	5%	1%	0.321	13%	3%
180	1	0.278	1%	0%	0.304	4%	0%	0.351	8%	0%	0.456	19%	0%
	2	0.275	0%	0%	0.296	2%	0%	0.352	8%	1%	0.454	18%	2%
	4	0.273	1%	0%	0.292	3%	0%	0.342	8%	1%	0.450	18%	4%
	5	0.271	1%	0%	0.293	3%	1%	0.340	8%	2%	0.447	18%	4%
	10	0.269	1%	0%	0.292	3%	1%	0.340	8%	2%	0.445	18%	5%
360	1	0.488	2%	0%	0.527	5%	0%	0.601	13%	0%	0.746	27%	0%
	2	0.485	1%	0%	0.520	5%	1%	0.598	13%	2%	0.744	27%	5%
	4	0.477	1%	0%	0.512	4%	1%	0.593	12%	3%	0.738	27%	8%
	5	0.474	1%	0%	0.513	5%	1%	0.592	13%	4%	0.733	27%	9%
	10	0.473	1%	0%	0.510	5%	1%	0.585	12%	4%	0.733	27%	10%
600	1	0.702	1%	0%	0.744	5%	0%	0.815	13%	0%	0.923	23%	0%
	2	0.700	2%	0%	0.738	6%	1%	0.816	13%	4%	0.921	24%	10%
	4	0.688	1%	0%	0.728	5%	2%	0.811	13%	5%	0.918	24%	13%
	5	0.689	1%	1%	0.728	5%	2%	0.808	13%	6%	0.918	24%	13%
	10	0.682	1%	0%	0.726	6%	2%	0.802	13%	6%	0.915	25%	14%

Table 6. Simulated power of group sequential RSS method, *Prss* , percentage power difference between RSS and SRS P_d , difference of average sample sizes between RSS and SRS as percentage of the total sample, eff, total sample size N , correlation with the concomitant variable ρ , number of interim analysis L , set size, $k = 2$, and alternative $H_a : \mu = 0.2$ under O'Brien-Fleming boundaries.

		ρ											
		0.25			0.5			0.75			1		
N	L	Prss	Pd	eff	Prss	Pd	eff	Prss	Pd	eff	Prss	Pd	eff
60	1	0.358	1%	0%	0.373	3%	0%	0.406	6%	0%	0.467	12%	0%
	2	0.350	0%	0%	0.369	2%	0%	0.403	6%	1%	0.474	13%	1%
	4	0.351	1%	0%	0.368	2%	1%	0.405	6%	1%	0.466	12%	3%
	5	0.351	1%	0%	0.371	3%	1%	0.402	6%	1%	0.467	12%	3%
	10	0.351	1%	0%	0.370	3%	1%	0.403	6%	2%	0.469	12%	4%
120	1	0.599	1%	0%	0.626	3%	0%	0.676	8%	0%	0.755	16%	0%
	2	0.598	1%	0%	0.626	3%	1%	0.674	8%	2%	0.754	16%	4%
	4	0.592	1%	0%	0.620	4%	1%	0.669	8%	3%	0.746	16%	6%
	5	0.593	1%	0%	0.616	3%	1%	0.669	9%	3%	0.746	16%	6%
	10	0.587	1%	0%	0.613	4%	1%	0.663	9%	3%	0.744	17%	7%
180	1	0.775	1%	0%	0.796	3%	0%	0.845	8%	0%	0.902	14%	0%
	2	0.770	1%	0%	0.795	3%	1%	0.843	8%	3%	0.899	14%	6%
	4	0.765	1%	0%	0.789	3%	1%	0.831	7%	4%	0.895	14%	8%
	5	0.763	1%	0%	0.785	3%	1%	0.829	7%	4%	0.893	14%	8%
	10	0.759	1%	1%	0.783	3%	2%	0.828	8%	5%	0.891	14%	9%
360	1	0.970	0%	0%	0.977	1%	0%	0.987	2%	0%	0.996	3%	0%
	2	0.969	0%	0%	0.977	1%	2%	0.987	2%	6%	0.995	3%	11%
	4	0.967	0%	0%	0.974	1%	2%	0.986	2%	5%	0.995	3%	9%
	5	0.967	0%	0%	0.975	1%	2%	0.985	2%	5%	0.995	3%	9%
	10	0.963	0%	0%	0.973	1%	2%	0.985	2%	5%	0.995	3%	10%
600	1	0.999	0%	0%	0.999	0%	0%	1.000	0%	0%	1.000	0%	0%
	2	0.999	0%	0%	0.999	0%	2%	1.000	0%	5%	1.000	0%	8%
	4	0.999	0%	0%	0.999	0%	2%	1.000	0%	4%	1.000	0%	7%
	5	0.998	0%	0%	0.999	0%	2%	1.000	0%	4%	1.000	0%	8%
	10	0.998	0%	0%	0.999	0%	2%	1.000	0%	4%	1.000	0%	8%

Table 7. Simulated power of group sequential RSS method, *Prss* , percentage power difference between RSS and SRS P_d , difference of average sample sizes between RSS and SRS as percentage of the total sample, eff, total sample size N , correlation with the concomitant variable ρ , number of interim analysis L , set size, $k = 3$, and alternative $H_a : \mu = 0.2$ under O'Brien-Fleming boundaries.

		ρ											
		0.25			0.5			0.75			1		
N	L	Prss	Pd	eff	Prss	Pd	eff	Prss	Pd	eff	Prss	Pd	eff
60	1	0.355	1%	0%	0.382	4%	0%	0.447	10%	0%	0.576	23%	0%
	2	0.355	1%	0%	0.384	4%	1%	0.445	10%	1%	0.571	22%	3%
	4	0.356	1%	0%	0.383	4%	1%	0.443	10%	2%	0.572	23%	6%
	5	0.352	1%	0%	0.383	4%	1%	0.441	10%	2%	0.575	23%	6%
	10	0.359	1%	1%	0.384	4%	1%	0.440	10%	3%	0.571	23%	8%
120	1	0.605	1%	0%	0.649	6%	0%	0.727	13%	0%	0.857	26%	0%
	2	0.604	1%	0%	0.642	5%	1%	0.725	13%	3%	0.855	26%	8%
	4	0.595	1%	0%	0.638	5%	2%	0.717	13%	5%	0.850	27%	11%
	5	0.598	1%	0%	0.635	5%	2%	0.715	13%	5%	0.847	26%	11%
	10	0.590	2%	1%	0.631	6%	2%	0.713	14%	6%	0.848	27%	13%
180	1	0.778	2%	0%	0.814	5%	0%	0.878	12%	0%	0.960	20%	0%
	2	0.772	1%	0%	0.814	5%	2%	0.878	12%	5%	0.958	20%	12%
	4	0.766	1%	0%	0.808	5%	2%	0.871	11%	6%	0.957	20%	14%
	5	0.763	1%	1%	0.806	5%	3%	0.873	12%	7%	0.954	20%	15%
	10	0.761	1%	1%	0.803	5%	3%	0.864	11%	7%	0.956	21%	16%
360	1	0.971	0%	0%	0.982	2%	0%	0.993	3%	0%	1.000	3%	0%
	2	0.970	0%	1%	0.982	2%	3%	0.993	3%	9%	1.000	3%	18%
	4	0.968	1%	1%	0.979	2%	3%	0.992	3%	8%	0.999	4%	15%
	5	0.966	0%	1%	0.980	2%	3%	0.992	3%	8%	0.999	4%	15%
	10	0.966	1%	1%	0.978	2%	3%	0.992	3%	8%	0.999	4%	16%
600	1	0.999	0%	0%	0.999	0%	0%	1.000	0%	0%	1.000	0%	0%
	2	0.999	0%	1%	1.000	0%	3%	1.000	0%	7%	1.000	0%	11%
	4	0.999	0%	0%	0.999	0%	2%	1.000	0%	6%	1.000	0%	12%
	5	0.998	0%	1%	0.999	0%	3%	1.000	0%	6%	1.000	0%	12%
	10	0.999	0%	1%	0.999	0%	3%	1.000	0%	7%	1.000	0%	13%

Table 8**.** Monitoring of the ACGT 320 trial by using equally spaced O'Brien-Fleming designs with L=4 analyses and $\alpha = 0.05$**, and power of 85%.**

Time	Boundaries	Z-Statistic(sample size/treatment)		
		RSS(k=2)	RSS(k=3)	SRS
1	4.048	1.05(85)	3.97(75)	2.61(105)
2	2.862	3.33(170)	3.36(150)	3.452(210)
3	2.337			
4	2.024			