Data Science in Kyoto Women's University 1 (2026) 93 − 106

Developing Group Sequential Procedure to Multiple-Sample Settings Using Pairwise Comparisons

Rin Tomioka*, Reina Hara* and Tomohiro Nakamura*

Key Words and Phrases : pairwise comparisons among several populations; repeated confidence boundaries; type I familywise error rate; all-pairs power

1. Introduction

The group sequential procedure proposed by Pocock (1977) is often used to test for differences in the effects of two treatments in clinical trials of long-term analysis. This group sequential procedure is known as a method that reduces the average sample size because it allows conclusions at an early stage. In addition, methods for determining repeated confidence boundaries for tests based on a prespecified significance level include those proposed by O'Brien and Fleming (1979), DeMets and Ware (1980), Lan and DeMets (1983), Whitehead (1983) and Fleming Harrington and O'Brien (1984). Since this procedure involves repeated testing, it raises the issue of multiplicity due to inflated significance levels; however, this problem can be resolved by setting repeated confidence boundaries as proposed by Armitage (1969). Although group sequential procedures are typically used to test the difference in effects between two treatments, they have not been fully extended to sequential pairwise comparisons in multiple-sample settings with three or more groups.

In this study, we extend the group sequential procedure to allow for pairwise comparisons in the multiple-sample problem, based on the concept of multiple comparison procedure introduced by Tukey (1953). We derive an integration formula that enables stepwise calculation of complex integrals based on multivariate normal distributions by using the conditional distribution at each stage. Since the proposed group sequential testing framework is designed for simultaneous hypotheses, we adopt the all-pairs power defined by Ramsey (1978) as the power of the test, which represents the probability of detecting all true differences among the population means. Finally, under a pre-specified significance level, we conduct simulation studies comparing the repeated confidence boundaries and the required sample size based on Pocock (1977), O'Brien and Fleming (1979) and Lan and DeMets (1983) in the settings of several treatments with stages.

* Faculty of Data Science, Kyoto Women's University, Japan
Corresponding Author: Tomohiro Nakamura E-mail : nakamurat@kyoto-wu.ac.jp

2. Pairwise Comparisons in the Multiple-Sample Problem

The group sequential test is a procedure that sequentially evaluates the difference in effects between two treatments based on the group observations obtained at each stage. Each time point at which a test is performed based on the collected group observations is referred to as a stage, and the maximum number of tests, *K*, is predetermined. Here, we denote the number of treatments by $L(L \geq 3)$, and represent the response of each treatment by the random variables $X_1, \cdots, X_L$. In this setting, $X_1, \cdots, X_L$ are assumed to follow normal distributions $N(\mu_1, \sigma^2), \cdots, N(\mu_L, \sigma^2)$ respectively, where $\mu_1, \cdots, \mu_L$ represent the population means, σ^2 denotes the known population variance. We denote this by

$$X_l \sim N(\mu_l, \sigma^2), \quad l = 1, \cdots, L. \tag{2.1}$$

Table 2.1 Group of observations at each stage

stage	group of observations			
1	$x_{1,1}$	$x_{1,2}$	$\cdots$	x_{1,n_1}
	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$x_{L,1}$	$x_{L,2}$	$\cdots$	x_{L,n_1}
$\vdots$		$\vdots$		
k	$x_{1,n_{k-1}+1}$	$x_{1,n_{k-1}+2}$	$\cdots$	x_{1,n_k}
	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$x_{L,n_{k-1}+1}$	$x_{L,n_{k-1}+2}$	$\cdots$	x_{L,n_k}
$\vdots$		$\vdots$		
K	$x_{1,n_{K-1}+1}$	$x_{n_{1,K-1}+2}$	$\cdots$	x_{1,n_K}
	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$x_{L,n_{K-1}+1}$	$x_{L,n_{K-1}+2}$	$\cdots$	x_{L,n_K}

If we let g denote the sample size drawn at each stage, then the cumulative sample size n_k obtained for each treatment up to the $k(k = 1, \cdots, K)$ th stage is given by $n_k = gk$.

Table 2.1 shows the group observations obtained from 1st stage through *K*th stage. In this study, we perform pairwise comparisons of the effects for any two treatments $l, m(1 \leq l < m \leq L)$ among the L treatments. Thus simultaneous hypotheses to sequential testing at each stage are as follows :

$$\begin{cases} \text{Null hypotheses} & H_{12}^{(0)} : \mu_2 = \mu_1, \ \cdots, \ H_{lm}^{(0)} : \mu_m = \mu_l, \ \cdots, \ H_{L-1,L}^{(0)} : \mu_L = \mu_{L-1}, \\ \text{Alternative hypotheses} & H_{12}^{(1)} : \mu_2 \neq \mu_1, \ \cdots, \ H_{lm}^{(1)} : \mu_m \neq \mu_l, \ \cdots, \ H_{L-1,L}^{(1)} : \mu_L \neq \mu_{L-1}. \end{cases} \tag{2.2}$$

The sample mean $\bar{X}_{l,k}$ of treatment $l(l = 1, \cdots, L)$ at $k(k = 1, \cdots, K)$ th stage is represented by

$$\bar{X}_{l,k} = \sum_{j=n_{k-1}+1}^{n_k} \frac{X_{l,j}}{g} \sim N\left(\mu_l, \frac{\sigma^2}{g}\right), \ l = 1, \cdots, L, \ k = 1, \cdots, K, \tag{2.3}$$

then

$$Z_{l,k} = \frac{\sqrt{g}\bar{X}_{l,k}}{\sigma} \sim N\left(\frac{\sqrt{g}\mu_l}{\sigma}, 1\right),\ l = 1, \cdots, L,\ k = 1, \cdots, K. \tag{2.4}$$

In this case,

$$Z_{m,k} - Z_{l,k} \sim N\left(\frac{\sqrt{g}(\mu_m - \mu_l)}{\sigma}, 2\right),\ 1 \le l < m \le L,\ k = 1, \cdots, K, \tag{2.5}$$

and

$$U_{l,k} = \sum_{i=1}^{k} Z_{l,i} \sim N\left(\frac{k\sqrt{g}\mu_l}{\sigma}, k\right),\ l = 1, \cdots, L,\ k = 1, \cdots, K. \tag{2.6}$$

The test statistic at $k(k = 1, \cdots, K)$ th stage can be expressed as follows.

$$S_{(lm),k} = U_{m,k} - U_{l,k} = \sum_{i=1}^{k} (Z_{m,i} - Z_{l,i}) \sim N\left(\frac{k\sqrt{g}(\mu_m - \mu_l)}{\sigma}, 2k\right),\ 1 \le l < m \le L,\ k = 1, \cdots, K. \tag{2.7}$$

2.1. Determination of Repeated Confidence Boundaries

Conventionally, group sequential procedures that have been developed are characterized by repeated confidence boundaries, and their foundation is based on the theory proposed by Armitage (1969), which treated the problem as one of stochastic processes.

In this study, since simultaneous hypotheses are further addressed at each stage, we adopt the type I familywise error rate, which is the probability that at least one true null hypothesis is incorrectly rejected in multiple comparison procedures. Under the assumption that all ${}_LC_2$ null hypotheses $H_{lm}^{(0)} : \mu_m = \mu_l (1 \le l < m \le L)$ hold, the repeated confidence boundaries $b_1, b_2, \cdots, b_K$ for the test statistics $\{S_{(12),1}, \cdots, S_{(L-1,L),1}\}$ at 1st stage, $\{S_{(12),2}, \cdots, S_{(L-1,L),2}\}$ at 2nd stage, $\cdots$, and $\{S_{(12),K}, \cdots, S_{(L-1,L),K}\}$ at *K*th stage are determined as follows :

$$\begin{aligned}
\alpha_1^* &= \Pr(|S_{(12),1}| > b_1 \text{ or } \cdots \text{ or } |S_{(L-1,L),1}| > b_1),\\
\alpha_2^* &= \Pr(|S_{(12),1}| \le b_1, \cdots, |S_{(L-1,L),1}| \le b_1, \{|S_{(12),2}| > b_2 \text{ or } \cdots \text{ or } |S_{(L-1,L),2}| > b_2\}),\\
&\vdots\\
\alpha_k^* &= \Pr(|S_{(12),1}| \le b_1, \cdots, |S_{(L-1,L),1}| \le b_1, \cdots, |S_{(12),k-1}| \le b_{k-1}, \cdots, |S_{(L-1,L),k-1}| \le b_{k-1},\\
&\qquad \{|S_{(12),k}| > b_k \text{ or } \cdots \text{ or } |S_{(L-1,L),k}| > b_k\}),\\
&\vdots\\
\alpha_K^* &= \Pr(|S_{(12),1}| \le b_1, \cdots, |S_{(L-1,L),1}| \le b_1, \cdots, |S_{(12),K-1}| \le b_{K-1}, \cdots, |S_{(L-1,L),K-1}| \le b_{K-1},\\
&\qquad \{|S_{(12),K}| > b_K \text{ or } \cdots \text{ or } |S_{(L-1,L),K}| > b_K\}).
\end{aligned} \tag{2.8}$$

Under the above condition holds, for a specified significance level α, we sequentially determine the repeated confidence boundaries $b_1, \cdots, b_K$ so as to satisfy

$$\alpha = \sum_{i=1}^{K} \alpha_i^*. \tag{2.9}$$

Determining the repeated confidence boundary b_k in equation (2.8) requires deriving the joint distribution of $S_{(lm),k}(1 \le l < m \le L, k = 1, \cdots, K)$, namely the probability density function (p.d.f.) of a $(K \times {}_LC_2)$-dimensional normal distribution, and performing multiple integrations. However, these multi-dimensional integrations, which depend on the number of treatments L and stages K, are not practically feasible. In this study, we derive the repeated integration formulae for computing α_k^* at each stage and apply it sequentially.

Here, under the assumption that all ${}_LC_2$ null hypotheses $H_{lm}^{(0)} : \mu_m = \mu_l (1 \le l < m \le L)$ hold, we have $\mu_1 = \cdots = \mu_L$. Furthermore, when determining repeated confidence boundaries, it is without loss of generality to set $\mu_1 = \cdots = \mu_L = 0$. At this point, equations (2.4) and (2.6) become as follows :

$$Z_{l,k} = \frac{\sqrt{g}\bar{X}_{l,k}}{\sigma} \sim N\left(\frac{\sqrt{g}\mu_l}{\sigma}, 1\right) \equiv N(0,1),\ l = 1, \cdots, L,\ k = 1, \cdots, K, \tag{2.10}$$

$$U_{l,k} = \sum_{i=1}^{k} Z_{l,i} \sim N\left(\frac{k\sqrt{g}\mu_l}{\sigma}, k\right) \equiv N(0,k),\ l = 1, \cdots, L,\ k = 1, \cdots, K. \tag{2.11}$$

1st stage

From (2.8), we have

$$\alpha_1^* = \Pr(|S_{(12),1}| > b_1 \text{ or } \cdots \text{ or } |S_{(L-1,L),1}| > b_1).$$

This expression can be rewritten as follows :

$$\begin{aligned}
1 - \alpha_1^* &= \Pr(|S_{(12),1}| \le b_1, \cdots, |S_{(L-1,L),1}| \le b_1) = \Pr(|Z_{2,1} - Z_{1,1}| \le b_1, \cdots, |Z_{L,1} - Z_{L-1,1}| \le b_1) \\
&= \Pr\left(\max_{\substack{l,m=1,\cdots,L \\ l \ne m}} |Z_{l,1} - Z_{m,1}| \le b_1\right) \\
&= \Pr(0 \le Z_{1,1} - Z_{2,1} \le b_1, \cdots, 0 \le Z_{1,1} - Z_{L,1} \le b_1, \max_{l=1,\cdots,L} Z_{l,1} = Z_{1,1}) \\
&\quad + \cdots + \Pr(0 \le Z_{L,1} - Z_{1,1} \le b_1, \cdots, 0 \le Z_{L,1} - Z_{L-1,1} \le b_1, \max_{l=1,\cdots,L} Z_{l,1} = Z_{L,1}) \\
&= \int_{-\infty}^{\infty} \Pr(0 \le z_{1,1} - Z_{2,1} \le b_1, \cdots, 0 \le z_{1,1} - Z_{L,1} \le b_1 | Z_{1,1} = z_{1,1}) g(z_{1,1}) dz_{1,1} \\
&\quad + \cdots + \int_{-\infty}^{\infty} \Pr(0 \le z_{L,1} - Z_{1,1} \le b_1, \cdots, 0 \le z_{L,1} - Z_{L-1,1} \le b_1 | Z_{L,1} = z_{L,1}) g(z_{L,1}) dz_{L,1} \\
&= L \int_{-\infty}^{\infty} \left(\int_{-b_1 + \tilde{z}_1}^{\tilde{z}_1} f(z_1) dz_1 \right)^{L-1} g(\tilde{z}_1) d\tilde{z}_1.
\end{aligned} \tag{2.12}$$

Here, both $f(z_{l,1})$ and $g(z_{l,1})$ are p.d.f. of $Z_{l,1} \sim N(0,1) (l = 1, \cdots, L)$, and $z_{l,1} \equiv z_1 \equiv \tilde{z}_1 (l = 1, \cdots, L)$.

$k(k = 2, \cdots, K)$ th stage

From (2.8), we have

$$\begin{aligned}\alpha_k^* = \Pr(&|S_{(12)\,,1}| \le b_1, \cdots, |S_{(L-1,L),1}| \le b_1, \cdots, |S_{(12)\,,k-1}| \le b_{k-1}, \cdots, |S_{(L-1,L),k-1}| \le b_{k-1},\\ &\{|S_{(12)\,,k}| > b_k \text{ or } \cdots \text{ or } |S_{(L-1,L),k}| > b_k\}).\end{aligned}$$

This expression can be rewritten as follows :

$$\begin{aligned}
&1 - (\alpha_1^* + \cdots + \alpha_k^*)\\
&= \Pr(|S_{(12)\,,1}| \le b_1, \cdots, |S_{(L-1,L),1}| \le b_1, \cdots, |S_{(12)\,,k-1}| \le b_{k-1}, \cdots, |S_{(L-1,L),k-1}| \le b_{k-1},\\
&\quad |S_{(12)\,,k}| \le b_k, \cdots, |S_{(L-1,L),k}| \le b_k)\\
&= \Pr[-b_1 \le Z_{2,1} - Z_{1,1} \le b_1, \cdots, -b_1 \le Z_{L,1} - Z_{1,1} \le b_1, \cdots,\\
&\quad -b_{k-1} \le Z_{2,1} + \cdots + Z_{2,k-1} - (Z_{1,1} + \cdots + Z_{1,k-1}) \le b_{k-1}, \cdots,\\
&\quad\ -b_{k-1} \le Z_{L,1} + \cdots + Z_{L,k-1} - (Z_{1,1} + \cdots + Z_{1,k-1}) \le b_{k-1},\\
&\quad\quad 0 \le Z_{1,1} + \cdots + Z_{1,k} - (Z_{2,1} + \cdots + Z_{2,k}) \le b_k, \cdots,\\
&\quad\quad\ 0 \le Z_{1,1} + \cdots + Z_{1,k} - (Z_{L,1} + \cdots + Z_{L,k}) \le b_k, \max_{l=1,\cdots,L}(Z_{l,1} + \cdots + Z_{l,k}) = (Z_{1,1} + \cdots + Z_{1,k})]\\
&\quad + \cdots + \Pr[-b_1 \le Z_{L,1} - Z_{1,1} \le b_1, \cdots, -b_1 \le Z_{L,1} - Z_{L-1,1} \le b_1, \cdots,\\
&\quad\ -b_{k-1} \le Z_{L,1} + \cdots + Z_{L,k-1} - (Z_{1,1} + \cdots + Z_{1,k-1}) \le b_{k-1}, \cdots,\\
&\quad\quad -b_{k-1} \le Z_{L,1} + \cdots + Z_{L,k-1} - (Z_{L-1,1} + \cdots + Z_{L-1,k-1}) \le b_{k-1},\\
&\quad\quad\ 0 \le Z_{L,1} + \cdots + Z_{L,k} - (Z_{1,1} + \cdots + Z_{1,k}) \le b_k, \cdots,\\
&\quad\quad\quad 0 \le Z_{L,1} + \cdots + Z_{L,k} - (Z_{L-1,1} + \cdots + Z_{L-1,k}) \le b_k, \max_{l=1,\cdots,L}(Z_{l,1} + \cdots + Z_{l,k}) = (Z_{L,1} + \cdots + Z_{L,k})]\\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \Pr[-b_1 \le Z_{2,1} - z_{1,1} \le b_1, \cdots, -b_1 \le Z_{L,1} - z_{1,1} \le b_1, \cdots,\\
&\quad -b_{k-1} \le Z_{2,1} + \cdots + Z_{2,k-1} - (z_{1,1} + \cdots + z_{1,k-1}) \le b_{k-1}, \cdots,\\
&\quad\ -b_{k-1} \le Z_{L,1} + \cdots + Z_{L,k-1} - (z_{1,1} + \cdots + z_{1,k-1}) \le b_{k-1},\\
&\quad\quad 0 \le z_{1,1} + \cdots + z_{1,k} - (Z_{2,1} + \cdots + Z_{2,k}) \le b_k, \cdots,\\
&\quad\quad\ 0 \le z_{1,1} + \cdots + z_{1,k} - (Z_{L,1} + \cdots + Z_{L,k}) \le b_k\\
&\quad\quad\quad |Z_{1,1}, = z_{1,1} \cdots, Z_{1,k} = z_{1,k}]g(z_{1,1}) \cdots g(z_{1,k-1})g(z_{1,k})dz_{1,1} \cdots dz_{1,k-1}dz_{1,k}\\
&\quad + \cdots + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \Pr[-b_1 \le z_{L,1} - Z_{1,1} \le b_1, \cdots, -b_1 \le z_{L,1} - Z_{L-1,1} \le b_1, \cdots,\\
&\quad\ -b_{k-1} \le z_{L,1} + \cdots + z_{L,k-1} - (Z_{1,1} + \cdots + Z_{1,k-1}) \le b_{k-1}, \cdots,\\
&\quad\quad -b_{k-1} \le z_{L,1} + \cdots + z_{L,k-1} - (Z_{L-1,1} + \cdots + Z_{L-1,k-1}) \le b_{k-1},\\
&\quad\quad\ 0 \le z_{L,1} + \cdots + z_{L,k} - (Z_{1,1} + \cdots + Z_{1,k}) \le b_k, \cdots,\\
&\quad\quad\quad 0 \le z_{L,1} + \cdots + z_{L,k} - (Z_{L-1,1} + \cdots + Z_{L-1,k}) \le b_k\\
&\quad\quad\quad |Z_{L,1} = z_{L,1}, \cdots, Z_{L,k} = z_{L,k}]g(z_{L,1}) \cdots g(z_{L,k-1})g(z_{L,k})dz_{L,1} \cdots dz_{L,k-1}dz_{L,k}\\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\left(\int_{-b_k+\sum_{i=1}^{k} z_{1,i}}^{\sum_{i=1}^{k} z_{1,i}}\int_{-b_{k-1}+\sum_{i=1}^{k-1} z_{1,i}}^{b_{k-1}+\sum_{i=1}^{k-1} z_{1,i}} f(u_{2,k-1})\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{(z_{2,k} - u_{2,k-1})^2}{2}\right\}du_{2,k-1}dz_{2,k}\right)\\
&\quad \times\cdots\times\left(\int_{-b_k+\sum_{i=1}^{k} z_{1,i}}^{\sum_{i=1}^{k} z_{1,i}}\int_{-b_{k-1}+\sum_{i=1}^{k-1} z_{1,i}}^{b_{k-1}+\sum_{i=1}^{k-1} z_{1,i}} f(u_{L,k-1})\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{(z_{L,k} - u_{L,k-1})^2}{2}\right\}du_{L,k-1}dz_{L,k}\right)\\
&\quad\ \times g(z_{1,1}) \cdots g(z_{1,k-1})g(z_{1,k})dz_{1,1} \cdots dz_{1,k-1}dz_{1,k}\\
&\quad + \cdots + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\left(\int_{-b_k+\sum_{i=1}^{k} z_{L,i}}^{\sum_{i=1}^{k} z_{L,i}}\int_{-b_{k-1}+\sum_{i=1}^{k-1} z_{L,i}}^{b_{k-1}+\sum_{i=1}^{k-1} z_{L,i}} f(u_{1,k-1})\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{(z_{1,k} - u_{1,k-1})^2}{2}\right\}du_{1,k-1}dz_{1,k}\right)\\
&\quad \times\cdots\times\left(\int_{-b_k+\sum_{i=1}^{k} z_{L,i}}^{\sum_{i=1}^{k} z_{L,i}}\int_{-b_{k-1}+\sum_{i=1}^{k-1} z_{L,i}}^{b_{k-1}+\sum_{i=1}^{k-1} z_{L,i}} f(u_{L-1,k-1})\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{(z_{L-1,k} - u_{L-1,k-1})^2}{2}\right\}du_{L-1,k-1}dz_{L-1,k}\right)\\
&\quad\ \times g(z_{L,1}) \cdots g(z_{L,k-1})g(z_{L,k})dz_{L,1} \cdots dz_{L,k-1}dz_{L,k}
\end{aligned}$$

$$= L \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\int_{-b_k+\tilde{u}_{k-1}+\tilde{z}_k}^{\tilde{u}_{k-1}+\tilde{z}_k} \int_{-b_{k-1}+\tilde{u}_{k-1}}^{b_{k-1}+\tilde{u}_{k-1}} f(u_{k-1}) \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{(z_k - u_{k-1})^2}{2} \right\} du_{k-1} dz_k \right)^{L-1}$$
$$\times f(\tilde{u}_{k-1}) g(\tilde{z}_k) d\tilde{u}_{k-1} d\tilde{z}_k. \tag{2.13}$$

Here, $g(z_{l,k})$ and $f(u_{l,k-1})$ are the p.d.f. of $Z_{l,k} \sim N(0,1)(l = 1, \cdots, L, k = 1, \cdots, K)$ and $U_{l,k-1} = Z_{l,1} + \cdots + Z_{l,k-1} \sim N(0, k-1)(l = 1, \cdots, L-1, k = 2, \cdots, K)$, respectively. Moreover, we set $u_{l,k-1} \equiv u_{k-1} \equiv \tilde{u}_{k-1}(l = 1, \cdots, L-1, k = 2, \cdots, K)$ and $z_{l,k} \equiv z_k \equiv \tilde{z}_k(l = 1, \cdots, L, k = 1, \cdots, K)$.

In the group sequential procedure, given the repeated confidence boundaries $b_1, \cdots, b_K$ at each stage and the observed values $s_{(12),1}, \cdots, s_{(L-1,L),1}, \cdots, s_{(12),K}, \cdots, s_{(L-1,L),K}$ of the test statistics $S_{(12),1}, \cdots, S_{(L-1,L),1}, \cdots, S_{(12),K}, \cdots, S_{(L-1,L),K}$, the decision rules are as follows:

- $k(k = 1, \cdots, K-1)$th stage

(*a*) If $|s_{(lm),k}| \leq b_k$ for all $l, m(1 \leq l < m \leq L)$, then defer the decision and retain observation until the next stage.

(*b*) If $|s_{(lm),k}| > b_k$ for at least one pair $(l, m)(1 \leq l < m \leq L)$, then terminate the test based on the following decision.

○ If $|s_{(lm),k}| \leq b_k (1 \leq l < m \leq L)$, then accept the corresponding null hypothesis $H_{lm}^{(0)}$.

○ If $|s_{(lm),k}| > b_k (1 \leq l < m \leq L)$, then reject the corresponding null hypothesis $H_{lm}^{(0)}$.

- Kth stage

(*c*) If $|s_{(lm),K}| \leq b_K$ for all $l, m(1 \leq l < m \leq L)$, then accept the corresponding null hypotheses $H_{lm}^{(0)}$ and terminate the test.

(*d*) If $|s_{(lm),K}| > b_K$ for at least one pair $(l, m)(1 \leq l < m \leq L)$, then terminate the test based on the following decision.

○ If $|s_{(lm),K}| \leq b_K (1 \leq l < m \leq L)$, then accept the corresponding null hypothesis $H_{lm}^{(0)}$.

○ If $|s_{(lm),K}| > b_K (1 \leq l < m \leq L)$, then reject the corresponding null hypothesis $H_{lm}^{(0)}$.

2.2. α Spending Function and Repeated Confidence Boundaries

We describe three representative methods for determining repeated confidence boundaries.

(1) Pocock's method (1977)

This method uses a common value Z across all stages. The value of Z is determined by applying the Armitage's method using the repeated confidence boundary

$$b_k = Z \cdot \sqrt{k}, \quad k = 1, \cdots, K \tag{2.14}$$

to obtain α_k^*, and adjusting Z until the condition $\alpha = \sum_{k=1}^{K} \alpha_k^*$ is satisfied. That is, this method fixes Z at a constant value across all stages and determines repeated confidence boundaries so

that the overall significance level α specified in advance is maintained over a maximum of K tests. Such a method has a high likelihood of early termination, and if early stopping does not occur, the final analysis tends to be conservative.

(2) O'Brien and Fleming's method (1979)

This method uses $U\sqrt{K/k}$. The value of U is determined based on the repeated confidence boundaries

$$b_k = U\sqrt{\frac{K}{k}}\sqrt{k} = U\cdot\sqrt{K}, \quad k = 1, \cdots, K \tag{2.15}$$

and is obtained through Monte Carlo simulation.

(3) Lan and DeMets' method (1983)

This method does not require the number of interim analyses to be specified in advance; analyses can be scheduled at equal or unequal intervals, and confidence boundaries can be determined at any time. As a feature of this method, the following alpha spending functions are assumed to allocate the significance level across stages, and based on this allocation, the Armitage's method is applied to determine the repeated confidence boundaries. This method mainly consists of three approaches. These approaches are characterized by how they allocate the significance level α, and are referred to as Normal method, Log method and *αt* method.

(a) Normal method

In this method, let Φ denote the standard normal distribution and $Z_{\frac{\alpha}{2}}$ is the $\frac{\alpha}{2}\%$ point of the standard normal distribution. Then, the α spending function α_k is defined as follows :

$$\alpha_k = 2\left\{1 - \Phi\left(Z_{\frac{\alpha}{2}}\sqrt{\frac{K}{k}}\right)\right\}, \quad k = 1, \cdots, K. \tag{2.16}$$

This method is characterized by assigning smaller probabilities at early stages, with the probabilities increasing as the number of stages grows.

(b) Log method

In this method, the α spending function α_k is defined using the following logarithmic function.

$$\alpha_k = \alpha\log\left\{1 + (e-1)\frac{k}{K}\right\}, \quad k = 1, \cdots, K. \tag{2.17}$$

This method is characterized by assigning larger probabilities at early stages, with the probabilities decreasing as the number of stages increases.

(*c*) *αt* method

In this method, let $t = k/K$, and define the α spending function α_k using the following linear function.

$$\alpha_k = \alpha\times\frac{k}{K} = \alpha t, \quad k = 1, \cdots, K. \tag{2.18}$$

This approach maintains a constant probability α_k across all stages and subsequently determines b_k. By the above methods, the significance level α_k is calculated for each stage, and the incremental value $\alpha_k^* = \alpha_k - \alpha_{k-1}$ is then derived. This α_k^* satisfies the following relationship.

$$\alpha = \sum_{k=1}^{K} \alpha_k^*. \tag{2.19}$$

2.3. Power and Required Sample Size

In group sequential procedures for the difference between two population means, power is the probability of crossing the repeated confidence boundaries under the alternative hypothesis. Since this study involves group sequential procedure for simultaneous hypotheses with three or more groups, we adopt the overall all-pairs power, as defined by Ramsey (1978), which is the probability of detecting all pairs with differences in population means, as the power of test in this study.

Now, we let

$$\frac{\sqrt{g}\mu_l}{\sigma} = \Delta_l (l = 1, \cdots, L), \tag{2.20}$$

then (2.4) and (2.6) become as follows.

$$Z_{l,k} = \frac{\sqrt{g}\bar{X}_{l,k}}{\sigma} \sim N\left(\frac{\sqrt{g}\mu_l}{\sigma}, 1\right) = N(\Delta_l, 1),\ U_{l,k} = \sum_{i=1}^{k} Z_{l,i} \sim N\left(\frac{k\sqrt{g}\mu_l}{\sigma}, k\right) = N(k\Delta_l, k),\ l = 1, \cdots, L. \tag{2.21}$$

At each $k(k = 1, \cdots, K)$ th stage, the probability π_k^* of accepting all ${}_LC_2$ alternative hypotheses $H_{lm}^{(1)} : \mu_m \neq \mu_l (1 \leq l < m \leq L)$, where differences exist among population means, is expressed as follows:

$$\begin{aligned}
\pi_1^* &= \Pr(|S_{(12),1}| > b_1, \cdots, |S_{(L-1,L),1}| > b_1), \\
\pi_2^* &= \Pr(|S_{(12),1}| \leq b_1, \cdots, |S_{(L-1,L),1}| \leq b_1, |S_{(12),2}| > b_2, \cdots, |S_{(L-1,L),2}| > b_2), \\
&\vdots \\
\pi_k^* &= \Pr(|S_{(12),1}| \leq b_1, \cdots, |S_{(L-1,L),1}| \leq b_1, \cdots, |S_{(12),k-1}| \leq b_{k-1}, \cdots, |S_{(L-1,L),k-1}| \leq b_{k-1}, \\
&\qquad |S_{(12),k}| > b_k, \cdots, |S_{(L-1,L),k}| > b_k), \\
&\vdots \\
\pi_K^* &= \Pr(|S_{(12),1}| \leq b_1, \cdots, |S_{(L-1,L),1}| \leq b_1, \cdots, |S_{(12),K-1}| \leq b_{K-1}, \cdots, |S_{(L-1,L),K-1}| \leq b_{K-1}, \\
&\qquad |S_{(12),K}| > b_K, \cdots, |S_{(L-1,L),K}| > b_K).
\end{aligned} \tag{2.22}$$

In this case, the values of $\Delta_l = \sqrt{g}\mu_l/\sigma (l = 1, \cdots, L)$ are sequentially adjusted until the condition

$$1 - \beta = \pi_1^* + \cdots + \pi_K^* \tag{2.23}$$

is satisfied for the specified probability of a Type II error β. Since σ is known, the value of g

will be sequentially adjusted for the configuration of population means under the alternative hypotheses, $\mu_1, \cdots, \mu_L$. $\pi_k^*(k = 1, \cdots, K)$ is obtained by the following integrations.

1st stage

From (2.22), we have

$$\begin{aligned}
\pi_1^* &= \Pr(|S_{(12),1}| > b_1, \cdots, |S_{(L-1,L),1}| > b_1) = \Pr(|Z_{2,1} - Z_{1,1}| > b_1, \cdots, |Z_{L,1} - Z_{L-1,1}| > b_1) \\
&= \Pr\left(\min_{\substack{l,m=1,\cdots,L \\ l \neq m}} |Z_{l,1} - Z_{m,1}| > b_1 \right) \\
&= \Pr(Z_{2,1} - Z_{1,1} > b_1, \cdots, Z_{L,1} - Z_{1,1} > b_1, \min_{l=1,\cdots,L} Z_{l,1} = Z_{1,1}) \\
&\quad + \cdots + \Pr(Z_{1,1} - Z_{L,1} > b_1, \cdots, Z_{L-1,1} - Z_{L,1} > b_1, \min_{l=1,\cdots,L} Z_{l,1} = Z_{L,1}) \\
&= \int_{-\infty}^{\infty} \Pr(Z_{2,1} - z_{1,1} > b_1, \cdots, Z_{L,1} - z_{1,1} > b_1 | Z_{1,1} = z_{1,1}) g(z_{1,1}) dz_{1,1} \\
&\quad + \cdots + \int_{-\infty}^{\infty} \Pr(Z_{1,1} - z_{L,1} > b_1, \cdots, Z_{L-1,1} - z_{L,1} > b_1 | Z_{L,1} = z_{L,1}) g(z_{L,1}) dz_{L,1} \\
&= \sum_{l=1}^{L} \int_{-\infty}^{\infty} \prod_{\substack{m=1 \\ m \neq l}}^{L} \left(\int_{b_1 + z_{l,1}}^{\infty} f(z_{m,1}) dz_{m,1} \right) g(z_{l,1}) dz_{l,1}.
\end{aligned} \tag{2.24}$$

Here, both $f(z_{l,1}), g(z_{l,1})$ are p.d.f. of $Z_{l,1} \sim N(\Delta_l, 1)(l = 1, \cdots, L)$.

$k(k = 2, \cdots, K)$th stage

By (2.22),

$$\begin{aligned}
\pi_k^* &= \Pr(|S_{(12),1}| \leq b_1, \cdots, |S_{(L-1,L),1}| \leq b_1, \cdots, |S_{(12),k-1}| \leq b_{k-1}, \cdots, |S_{(L-1,L),k-1}| \leq b_{k-1}, \\
&\qquad |S_{(12),k}| > b_k, \cdots, |S_{(L-1,L),k}| > b_k) \\
&= \Pr[-b_1 \leq Z_{2,1} - Z_{1,1} \leq b_1, \cdots, -b_1 \leq Z_{L,1} - Z_{1,1} \leq b_1, \cdots, \\
&\quad -b_{k-1} \leq Z_{2,1} + \cdots + Z_{2,k-1} - (Z_{1,1} + \cdots + Z_{1,k-1}) \leq b_{k-1}, \cdots, \\
&\quad -b_{k-1} \leq Z_{L,1} + \cdots + Z_{L,k-1} - (Z_{1,1} + \cdots + Z_{1,k-1}) \leq b_{k-1}, \\
&\quad Z_{2,1} + \cdots + Z_{2,k} - (Z_{1,1} + \cdots + Z_{1,k}) > b_k, \cdots, \\
&\quad Z_{L,1} + \cdots + Z_{L,k} - (Z_{1,1} + \cdots + Z_{1,k}) > b_k, \min_{l=1,\cdots,L} (Z_{l,1} + \cdots + Z_{l,k}) = (Z_{1,1} + \cdots + Z_{1,k})] \\
&\quad + \cdots + \Pr[-b_1 \leq Z_{L,1} - Z_{1,1} \leq b_1, \cdots, -b_1 \leq Z_{L,1} - Z_{L-1,1} \leq b_1, \cdots, \\
&\quad -b_{k-1} \leq Z_{L,1} + \cdots + Z_{L,k-1} - (Z_{1,1} + \cdots + Z_{1,k-1}) \leq b_{k-1}, \cdots, \\
&\quad -b_{k-1} \leq Z_{L,1} + \cdots + Z_{L,k-1} - (Z_{L-1,1} + \cdots + Z_{L-1,k-1}) \leq b_{k-1}, \\
&\quad Z_{1,1} + \cdots + Z_{1,k} - (Z_{L,1} + \cdots + Z_{L,k}) > b_k, \cdots, \\
&\quad Z_{L-1,1} + \cdots + Z_{L-1,k} - (Z_{L,1} + \cdots + Z_{L,k}) > b_k, \min_{l=1,\cdots,L} (Z_{l,1} + \cdots + Z_{l,k}) = (Z_{L,1} + \cdots + Z_{L,k})] \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \Pr[-b_1 \leq Z_{2,1} - z_{1,1} \leq b_1, \cdots, -b_1 \leq Z_{L,1} - z_{1,1} \leq b_1, \cdots, \\
&\quad -b_{k-1} \leq Z_{2,1} + \cdots + Z_{2,k-1} - (z_{1,1} + \cdots + z_{1,k-1}) \leq b_{k-1}, \cdots, \\
&\quad -b_{k-1} \leq Z_{L,1} + \cdots + Z_{L,k-1} - (z_{1,1} + \cdots + z_{1,k-1}) \leq b_{k-1}, \\
&\quad Z_{2,1} + \cdots + Z_{2,k} - (z_{1,1} + \cdots + z_{1,k}) > b_k, \cdots, \\
&\quad Z_{L,1} + \cdots + Z_{L,k} - (z_{1,1} + \cdots + z_{1,k}) > b_k \\
&\quad |Z_{1,1}, = z_{1,1} \cdots, Z_{1,k} = z_{1,k}] g(z_{1,1}) \cdots g(z_{1,k-1}) g(z_{1,k}) dz_{1,1} \cdots dz_{1,k-1} dz_{1,k}
\end{aligned}$$

$$
\begin{aligned}
&+\cdots+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\Pr[-b_1\le z_{L,1}-Z_{1,1}\le b_1,\cdots,-b_1\le z_{L,1}-Z_{L-1,1}\le b_1,\cdots,\\
&\quad -b_{k-1}\le z_{L,1}+\cdots+z_{L,k-1}-(Z_{1,1}+\cdots+Z_{1,k-1})\le b_{k-1},\cdots,\\
&\quad\quad -b_{k-1}\le z_{L,1}+\cdots+z_{L,k-1}-(Z_{L-1,1}+\cdots+Z_{L-1,k-1})\le b_{k-1},\\
&\quad\quad\quad Z_{1,1}+\cdots+Z_{1,k}-(z_{L,1}+\cdots+z_{L,k})>b_k,\cdots,\\
&\quad\quad\quad\quad Z_{L-1,1}+\cdots+Z_{L-1,k}-(z_{L,1}+\cdots+z_{L,k})>b_k\\
&\quad\quad\quad\quad\quad |Z_{L,1}=z_{L,1},\cdots,Z_{L,k}=z_{L,k}]g(z_{L,1})\cdots g(z_{L,k-1})g(z_{L,k})dz_{L,1}\cdots dz_{L,k-1}dz_{L,k}\\
&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\left(\int_{b_k+\sum_{i=1}^{k}z_{1,i}}^{\infty}\int_{-b_{k-1}+\sum_{i=1}^{k-1}z_{1,i}}^{b_{k-1}+\sum_{i=1}^{k-1}z_{1,i}}f(u_{2,k-1})\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{(z_{2,k}-u_{2,k-1}-\Delta_2)^2}{2}\right\}du_{2,k-1}dz_{2,k}\right)\\
&\quad\times\cdots\times\left(\int_{b_k+\sum_{i=1}^{k}z_{1,i}}^{\infty}\int_{-b_{k-1}+\sum_{i=1}^{k-1}z_{1,i}}^{b_{k-1}+\sum_{i=1}^{k-1}z_{1,i}}f(u_{L,k-1})\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{(z_{L,k}-u_{L,k-1}-\Delta_L)^2}{2}\right\}du_{L,k-1}dz_{L,k}\right)\\
&\quad\quad\times g(z_{1,1})\cdots g(z_{1,k-1})g(z_{1,k})dz_{1,1}\cdots dz_{1,k-1}dz_{1,k}+\cdots\\
&+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\left(\int_{b_k+\sum_{i=1}^{k}z_{L,i}}^{\infty}\int_{-b_{k-1}+\sum_{i=1}^{k-1}z_{L,i}}^{b_{k-1}+\sum_{i=1}^{k-1}z_{L,i}}f(u_{1,k-1})\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{(z_{1,k}-u_{1,k-1}-\Delta_1)^2}{2}\right\}du_{1,k-1}dz_{1,k}\right)\\
&\quad\times\cdots\times\left(\int_{b_k+\sum_{i=1}^{k}z_{L,i}}^{\infty}\int_{-b_{k-1}+\sum_{i=1}^{k-1}z_{L,i}}^{b_{k-1}+\sum_{i=1}^{k-1}z_{L,i}}f(u_{L-1,k-1})\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{(z_{L-1,k}-u_{L-1,k-1}-\Delta_{L-1})^2}{2}\right\}du_{L-1,k-1}dz_{L-1,k}\right)\\
&\quad\quad\times g(z_{L,1})\cdots g(z_{L,k-1})g(z_{L,k})dz_{L,1}\cdots dz_{L,k-1}dz_{L,k}\\
&=\sum_{l=1}^{L}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\prod_{\substack{m=1\\m\neq l}}^{L}\left(\int_{b_k+u_{l,k-1}+z_{l,k}}^{\infty}\int_{-b_{k-1}+u_{l,k-1}}^{b_{k-1}+u_{l,k-1}}f(u_{m,k-1})\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{(z_{m,k}-u_{m,k-1}-\Delta_m)^2}{2}\right\}du_{m,k-1}dz_{m,k}\right)\\
&\quad\quad\times f(u_{l,k-1})g(z_{l,k})du_{l,k-1}dz_{l,k}\,. \qquad (2.25)
\end{aligned}
$$

Here, $g(z_{l,k}), f(u_{l,k-1})$ are the p.d.f. of $Z_{l,k}\sim N(\Delta_l,1)(l=1,\cdots,L,k=1,\cdots,K)$ and $U_{l,k-1}=Z_{l,1}+\cdots+Z_{l,k-1}\sim N((k-1)\Delta_l,k-1)(l=1,\cdots,L-1,k=2,\cdots,K)$ respectively. Since $\sqrt{g}\mu_l/\sigma=\Delta_l(l=1,\cdots,L)$ by (2.20), it follows that

$$
g=\frac{\sigma^2\Delta_l^2}{\mu_l^2},\ l=1,\cdots,L. \qquad (2.26)
$$

The maximum sample size per treatment is given by $n_{\max}=gK$.

3. Simulations

First, in the existing group sequential procedure for the difference between two population means with L =2, α =0.05, K =3, we compare the repeated confidence boundaries among Pocock's method, O'Brien and Fleming's method and Lan and DeMets' method in Table 3.1 and Fig. 3.1.

Table 3.1 Comparison among the repeated confidence boundaries (L =2, K = 3)

	b_1	b_2	b_3
Pocock′s method	2.298	3.250	3.980
O′Brien and Fleming′s method	3.471	3.471	3.471
Normal method	3.395	3.403	3.492
αt method	2.394	3.244	3.843
Log method	2.279	3.245	4.020

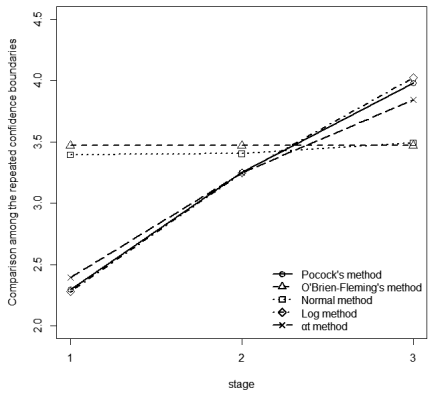

Fig. 3.1 Comparison among the repeated confidence boundaries (L = 2, K = 3)

O'Brien and Fleming's method takes the same value across all stages and shows a pattern almost identical to the Normal method. The other three methods exhibit increasing repeated confidence boundaries as the stages progress. Next, in this study, when α =0.05 and (L, K) = (3, 3), (4, 4), we compare the repeated confidence boundaries of Pocock's method, O'Brien and Fleming's method and Lan and DeMets' method.

Table 3.2 Comparison among the repeated confidence boundaries (L =3, K = 3)

	b_1	b_2	b_3
Pocock′s method	3.554	5.026	6.156
O′Brien and Fleming′s method	5.806	5.806	5.806
Normal method	5.202	5.514	5.868
αt method	3.882	5.276	5.963
Log method	3.732	5.259	5.972

Table 3.3 Comparison among the repeated confidence boundaries (L =4, K = 4)

	b_1	b_2	b_3	b_4
Pocock′s method	3.871	5.474	6.704	7.742
O′Brien and Fleming′s method	7.427	7.427	7.427	7.427
Normal method	6.124	6.580	7.043	7.528
αt method	4.305	5.888	6.774	7.680
Log method	4.143	5.797	6.698	7.748

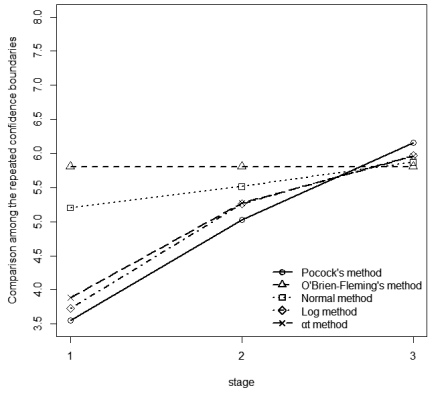

Fig. 3.2 Comparison among the repeated confidence boundaries ($L = 3$, $K = 3$)

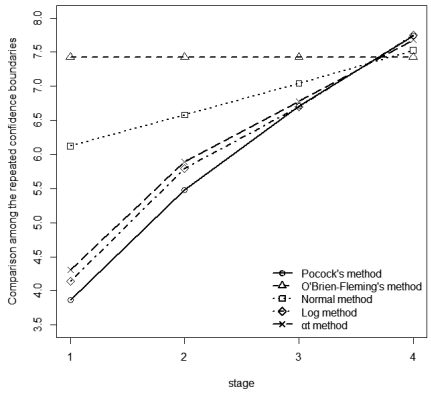

Fig. 3.3 Comparison among the repeated confidence boundaries ($L = 4$, $K = 4$)

Similar to the conventional group sequential procedure for the difference between two population means, the O'Brien and Fleming's method maintains the same value across all stages and exhibits a pattern almost identical to the Normal method. The other three methods show increasing repeated confidence boundaries as the stages progress. Moreover, as the numbers of L, K increase, the number of tests becomes larger, making the procedure more conservative; consequently, the values of the repeated confidence boundaries also increase with L, K.

Next, with α =0.05, σ^2 =1, and power $1-\beta$ =0.8, we compare the required sample size g at each stage for Pocock's method, O'Brien and Fleming's method and Lan and DeMets' method under the specified configurations of population means (μ_1, μ_2, μ_3) for K =3, L =3 and $(\mu_1, \mu_2, \mu_3, \mu_4)$ for K =4, L =4.

Table 3.4 Comparison of required sample size g at each stage (L =3, K =3, $1-\beta$ =0.8)

(μ_1, μ_2, μ_3)	$(1, 1.5, 2)$	$(1, 2, 3)$	$(1, 3, 5)$
Pocock′s method	91	23	6
O′Brien and Fleming′s method	196	49	13
Normal method	164	41	11
αt method	103	26	7
Log method	97	25	7

Table 3.5 Comparison of required sample size g at each stage (L =4, K =4, $1-\beta$ =0.8)

$(\mu_1, \mu_2, \mu_3, \mu_4)$	$(1, 1.5, 2, 2.5)$	$(1, 2, 3, 4)$	$(1, 3, 5, 7)$
Pocock′s method	103	26	7
O′Brien and Fleming′s method	298	75	19
Normal method	214	54	14
αt method	121	31	8
Log method	114	29	8

From the 1st stage, O'Brien and Fleming's method and Normal method, which have large repeated confidence boundaries, require a larger sample size g. Furthermore, as the numbers of L, K increase and the number of tests becomes larger, making the procedure more conservative; consequently, the required sample size g also increases with L, K.

4. Further discussions

This study extends the group sequential procedure, which addresses one or two sample problems, to multi-sample problems involving three or more sample problems. Furthermore, based on the concept of multiple comparison methods introduced by Tukey (1953), we attempt to develop a group sequential procedure that performs pairwise comparisons for comparing three or more population means. This study adopts the definition of overall power for multiple comparison methods proposed by Ramsey (1978), namely the all-pairs power. In simulation studies, we compare the aspects of repeated confidence boundaries and the required sample size among Pocock (1977), O'Brien and Fleming (1979) and Lan and DeMets (1983) in the settings of several treatments with stages.

However, we still have some improvements to consider. Although detecting all differences among population means is challenging, the any-pair power from the same multiple comparison framework can also be considered. We'd like to derive a similar integral formula for this and compare the two types of power in future. In addition, the acceptance of the null hypothesis in the group sequential procedure usually occurs only at the final stage. The test is terminated as soon as any hypothesis is rejected, which means that for hypotheses not rejected, no decision is made before the test ends. Therefore, we'd like to consider a decision rule that allows hypotheses not rejected to continue being tested in the next stage. In this case, since the rejected test statistics will not be tested in subsequent stages, the number of statistics decreases at each

stage, making it necessary to recalculate the repeated confidence boundaries. This approach will be expected to enable the development of a group sequential procedure with higher statistical power.

References

[1] Armitage, P., McPherson, C. K. and Rowe, B. C. (1969). Repeated Significance Tests on Accumulating data, J. R. Statist. Soc. A, 132, 235-244.

[2] DeMets, D. L. and Ware, J. H. (1980). Group Sequential Methods for Clinical Trials with a One-sided Hypothesis. Biometrika 67,3, 651-660.

[3] Fleming, T. R., Harrington, D. P. and O'Brien, P. C. (1984). Designs for Group Sequential Tests. Controlled Clinical Trials 5, 348-361.

[4] Lan, K. K. G. and DeMets, D. L. (1983). Discrete Sequential Boundaries for Clinical Trials. Biometrika, 70, 659-663.

[5] O'Brien, P. C. and Fleming, T. R. (1979). A Multiple Testing Procedure for Clinical Trials. Biometrics, 35, 549-556.

[6] Pocock, S. J. (1977). Group Sequential Methods in the Design and Analysis of Clinical Trials. Biometrika, 64, 191-199.

[7] Ramsey, P. H. (1978). Differences Between Pairwise Multiple Comparisons. Journal of the American Statistical Association, 73, 479-485.

[8] Tukey, J. W. (1953). The Problem of Multiple comparisons. Mimeographed monograph.

[9] Whitehead, J. (1983). The Design and Analysis of Sequential Clinical Trials. Chichester, Ellis Horwood.

Received: 19 February 2026