

## Research Article

# A sequential test procedure for monitoring a singular safety and efficacy outcome

**Alan D. Hutson**

University at Buffalo, Department of Statistics, Farber Hall Room 249A, 3435 Main St., Buffalo, NY 14214-3000.

## Abstract

*In this note we describe a modification of the sequential probability ratio test (SPRT) developed for the purpose of “flagging” a significant increase in the mortality rate of a treatment relative to a control while ensuring that double-blinding and the Type I error for the primary test of efficacy, also based on mortality rates, is not compromised.*

**Key words:** *Interim analysis; safety monitoring; sequential testing.*

---

Φ To whom correspondence should be addressed: *E-mail:* [ahutson@buffalo.edu](mailto:ahutson@buffalo.edu)

## Introduction

Drug trials go through different phases. In phase I trials the primary concern is safety, the subjects are typically healthy volunteer and patient studies, and the primary objective is to determine the maximum tolerated dose (MTD). This is followed by a phase II trial, which build upon the results of the phase I trial. The primary goal of a phase II trial is to determine the optimal method of administration and examine potential efficacy. If the phase II trial demonstrates that the drug may be reasonably safe and potentially effective a phase III trial may be carried forth. The primary goal of a phase III trial is to compare the effectiveness of the new treatment with that of existing treatments or placebo.

In a majority of phase III clinical trials subject measurements can be divided into distinct efficacy and safety variables. With the exception of truly sequential trials the primary efficacy variables are usually analyzed at a few well defined points in time (typically based upon calendar time or accrual milestones). In contrast, safety is monitored continuously through the generation of adverse event (AE) reports to the respective internal review board (IRB) and review by the principal investigator (PI), with summary data safety monitoring board (DSMB) reports generated at regular intervals, e.g. quarterly or bi-yearly. In a majority of high-risk clinical trials the DSMB will also monitor the flow of AE reports in real-time.

In the specific case of double-blind controlled trials of a new treatment versus standard of care, or placebo, how might one monitor safety "continuously" and efficacy at distinct points in time when the primary safety and efficacy variable is mortality. In addition, if the DSMB wishes to remain blinded to treatment assignment unless there is a true safety issue, what type of information is needed in order for them to make an informed decision in conjunction with real

time monitoring of AE reports? To generate a formal statistical rule to tackle this problem requires the principal investigators to design the trial around the efficacy outcome in the traditional sense, while the DSMB needs to determine what are the unacceptable differences between the new therapy and control in the opposite (unsafe) direction.

In this note we develop a one-sided safety monitoring rule based upon a modification of Wald's classical sequential probability ratio test (SPRT)<sup>1</sup>. This rule is to be used in conjunction with a randomized block design where the hypothesis involving efficacy is a simple comparison of two mortality proportions of the form  $H_0 : p_1 = p_2$  versus  $H_1 : p_1 < p_2$ . The general features of the safety monitoring procedure are as follows:

1. Mortality can be evaluated within a fixed time period shortly following treatment,
2. It is one-sided in the sense that a safety problem is not flagged if the new treatment appears efficacious,
3. Requires the input of members of the DSMB to determine what an unsafe rate might be for the new therapy relative to control given the design parameters of the trial,
4. Is continuously sequential in terms of blocks of subjects being the unit of time,
5. It is easily explained to non-statisticians through the use of examples,
6. Maintains the Type I error control for the test of efficacy conditioned upon the relative safety of the experimental therapy.

The motivation of this methodology stemmed from the request of our DSMB to generate a safety monitoring rule based upon statistical methods for the DCA-MALA clinical trial of malaria in Ghana<sup>2</sup>. This trial was designed as a randomized, double-blind, placebo-controlled single-center trial. The trial was originally designed to enroll n=1500 subjects, but was terminated earlier due to financial reasons. For this specific trial of dichloroacetate (DCA) versus placebo the

measure of efficacy was 28-day mortality, and the primary measure of safety was also 28-day mortality. Ultimately the DSMB wanted a simple statement after each block of subjects completed the trial: “remain blinded at this point in time” or that the “blind be broken at this point in time.” Unblinding the study is used to mean group summary statistics will be analyzed and presented in an unblinded fashion. This does not necessarily mean that data at the individual subject level is unblinded. Also note that the principal investigators would remain blinded even if the DSMB were to take an unblinded look at this trial. The use of this rule still allowed us to monitor secondary measures of safety in the standard way<sup>3,4</sup>. In addition, adverse events involving mortality were still monitored on a case-by-case basis (in a blinded manner).

### Statistical Methods

We will employ a version of a sequential test for the purpose of generating a safety monitoring rule, which will basically flag a problem for the DSMB with respect to a disproportionate amount of mortalities for a new therapy relative to the standard of care or placebo. Note that this rule does not terminate the clinical trial, it only suggests unblinding the trial for the purpose of more intense scrutiny. The sequential test is based upon a modification of the SPRT, initially developed by Wald<sup>1</sup>, and is a procedure for testing a simple null hypothesis versus a simple alternative hypothesis continuously in time. With respect to the new safety monitoring plan, “continuously in time” will refer to blocks of subjects who complete the trial as opposed to testing after individual subjects complete the trial. Therefore, in order to implement this rule effectively the design of the trial should be of the form of a randomized block design. Let  $N=n_1+n_2$  denote the total number of subjects for the new treatment plus the standard treatment,  $K$  denote the number of

blocks, and  $N_i=n_{1i} + n_{2i}$  denote the number of subjects in block  $i$ ,  $i=1, 2, \dots, K$ , where  $N_i=N/K$ .

For the purposes of our safety monitoring plan we set the null “efficacy” hypothesis to correspond to the original study design and the alternative “safety” hypothesis to correspond to unsafe rates of the new therapy relative to control. Let  $p_1$  and  $p_2$  denote the event rates of the new therapy and control group, respectively. Then the safety data monitoring rule (SDMR) consists of testing

$$\begin{aligned} H_0 &: \theta_0 = \text{efficacy}(p_1, p_2), \\ H_1 &: \theta_1 = \text{safety}(p_1, p_2), \end{aligned}$$

after blocks of  $N_i = n_{1i} + n_{2i}$  subjects complete the study. The DSMB chairperson is then notified of the results after each test is carried out per block. We strongly recommend that the values for  $p_1$  and  $p_2$  corresponding to  $H_0$  be chosen based upon the original study design pertaining to efficacy. For example, in the DCA-MALA trial the mortality rates from which the trial was designed were  $p_1 = 0.19$  and  $p_2 = 0.25$ , for DCA and placebo, respectively. For  $H_1$ , the “safety” hypothesis, the DSMB with the guidance of a simulation study deemed  $p_1 = 0.28$  to be an unacceptable death rate in the DCA arm given a placebo death rate of  $p_2 = 0.25$ , and accounting for statistical noise. The mathematical details for the efficacy and safety “functions” are contained in Appendix A. Through simulations we illustrated that if the placebo death rate is lower than anticipated the decision to recommend unblinding the trial will be earlier given the same relative differences in adverse event rates, e.g.  $p_1 = 0.18$  to  $p_2 = 0.15$  for DCA relative to placebo. Note that the unblinding rule can be easily modified to accommodate other types of outcomes such as mean differences.

**Test Statistic.** Denote the estimates of the proportions  $p_1$  and  $p_2$  as

$$\hat{p}_{1i} = \frac{\text{number of mortalities in the new therapy arm}}{n_{1i}}$$

$$\hat{p}_{2i} = \frac{\text{number of mortalities in the control arm}}{n_{2i}}$$

where  $i=1, 2, \dots, K$ , and again  $K$  denotes the number of blocks. The safety monitoring test statistic  $\lambda_i$  is then simply a function of  $\hat{p}_{1i}$  and  $\hat{p}_{2i}$  updated after every block  $i$ . For the DCA-MALA study we set  $B=75$  and  $n_{1i}=n_{2i}=10$  based upon efficacy considerations. Note that  $n_{1i}$  and  $n_{2i}$  have to be large enough to produce a meaningful value for  $\lambda_i$ . The details for the calculation of  $\lambda_i$  are contained in Appendix A. If  $\lambda_i \geq A$  then stop and reject  $H_0$  (recommend unblinding the study). If  $\lambda_i \leq B$  then stop and accept  $H_0$  (reset the monitoring rule), else continue to monitor the safety of the study. The consequences of resetting the monitoring rule are examined further in Section.

Classical statistical theory dictates that  $A \leq \frac{1-\alpha}{\beta}$  and  $B \geq \frac{\beta}{1-\alpha}$ . The crude stopping

bounds  $A = \frac{1-\alpha}{\beta}$  and  $B = \frac{\beta}{1-\alpha}$  work well

for our purpose. This is demonstrated via the the simulation study in Section 3. The parameters  $\alpha$  and  $\beta$  correspond roughly to traditional fixed sample size Type I and Type II error rates. Hence, we will typically choose  $\beta$  to be much smaller than  $\alpha$  if the primary goal is safety monitoring. For the DCA-MALA trial we determined that the appropriate levels would be to set  $\alpha=0.20$  and  $\beta=10^{-8}$  such that it would be unlikely that

the test terminates earlier and we accept  $H_0$ , yet the test would terminate quickly if there is a safety concern. Setting  $\alpha=0.20$  and  $\beta=10^{-8}$  corresponds to the stopping bounds of  $A=9.9999$  and  $B=0.00001$ . Therefore, if the test errs it will err on the conservative side in terms of unblinding the study early. The parameters  $\alpha$  and  $\beta$  may be adjusted by the DSMB in order to relax or tighten the monitoring rule during the course of the study.

Using the candidate choices of  $\alpha=0.20$  and  $\beta=10^{-8}$  (approximately null) in the DCA-MALA trial it was determined through simulation that  $\lambda_i$  would never be less than  $B$  prior to the first planned interim analysis at 50% accrual. If this scenario did occur it would have indicated that DCA is substantially more effective than originally anticipated and that the safety of DCA in terms of the mortality rates shouldn't be of concern to the EAC at that current point in time. Therefore, we propose that if  $\lambda_i < B$  during any point in the study that the SPRT decision rule be reset at block  $i-1$ , i.e. restart the safety monitoring rule one block back relative to the current point in time (block  $i$ ) as the new "time 0." This provides a one block "burn-in" period to reset the rule. Even after resetting the test statistic due to the outstanding performance of the new therapy, a short run of deaths could occur favoring control such that  $\lambda$  reverses its path and crosses  $A$ . This rare event (given a long-term past history of treatment efficacy) would not stop the trial, however, the recommendation to unblind the trial at that point in time would be made to the DSMB Chair. It would then have to be determined whether this run was due to chance alone, or some deterministic cause such as a bad batch of drug.

### Simulation Study

The following simulation study is used to illustrate the proportion of times out of 10,000 simulations that the decision to "remain blinded" or "unblind the study" would

occur over possible choices of  $\alpha=0.05, 0.10, 0.20$  and  $\beta=10^{-7}, 0.05, 0.10$  for a clinical trial with sample size fixed at  $n=1500$ , broken into  $K=75$  blocks. The decision to "reset the trial" the trial, as described above, is built-in to the simulation study. The numbers are similar to the operating characteristics of the DCA-MALA trial, however, there are no planned interim analyses. In addition, the median time that  $\lambda$  crosses  $A$  or  $B$  is given. For any specific trial of interest a similar simulation study should be undertaken in order to determine the appropriate parameter values for  $\alpha$  and  $\beta$ .

In this specific simulation study the efficacy hypothesis  $\theta_0(p_1, p_2)$  was fixed at  $\theta_0(0.20, 0.25)$  as determined by a hypothetical trial design. Assume that the DSMB decided the safety hypothesis should be  $\theta_1(0.30, 0.25)$ . The simulations were then carried out given different scenarios of "true" mortality rates  $(p_1, p_2)$ . The pairs  $(p_1, p_2)$  were set to  $(0.15, 0.25)$ ,  $(0.20, 0.25)$ ,  $(0.25, 0.25)$ ,  $(0.30, 0.25)$ ,  $(0.20, 0.15)$ , and  $(0.15, 0.10)$ , corresponding to the new treatment being more efficacious than planned, the new treatment being efficacious as planned, the new treatment being equivalent to placebo, the new treatment being worse than placebo at the correct placebo rate, the new treatment being worse than placebo at a lower placebo mortality rate, and the new treatment being worse than placebo at a very low and unexpected placebo mortality rate, respectively.

The simulation results are provided in Tables 1-9. Since we are primarily interested in testing for safety the choice of  $\alpha$  and  $\beta$  comes down to a tradeoff between stopping times and Type I error for this application. The column labeled "Median Sample Size" indicates the median time at which a decision rule is to be implemented if the true underlying mortality rates were  $p_1$  and  $p_2$ . If we were confident of the underlying truth with respect to  $p_1$  and  $p_2$  in terms of efficacy then it becomes a question of trade-offs.

**Table 1:** Simulation Results

$\alpha = .05$ and $\beta = 10^{-7}$				
$\theta_1(p_1, p_2)$		Median	Decision	
$p_1$	$p_2$	Sample Size	Unblind	Remain Blinded
0.15	0.25	1500	0.5%	95.5%
0.20	0.25	1500	6.3%	93.7%
0.25	0.25	1240	54.4.0%	45.6%
0.30	0.25	340	97.6%	2.4%
0.20	0.15	280	99.2%	0.8%
0.15	0.10	240	99.8%	0.2%

**Table 2:** Simulation Results

$\alpha = .1$ and $\beta = 10^{-7}$				
$\theta_1(p_1, p_2)$		Median	Decision	
$p_1$	$p_2$	Sample Size	Unblind	Remain Blinded
0.15	0.25	1500	1.3%	98.7%
0.20	0.25	1500	11.2%	88.8%
0.25	0.25	800	63.6%	36.4%
0.30	0.25	260	98.5%	1.5%
0.20	0.15	200	99.5%	0.5%
0.15	0.10	180	99.9%	0.1%

**Table 3:** Simulation Results

$\alpha = .2$ and $\beta = 10^{-7}$				
$\theta_1(p_1, p_2)$		Median	Decision	
$p_1$	$p_2$	Sample Size	Unblind	Remain Blinded
0.15	0.25	1500	3.6%	96.4%
0.20	0.25	1500	19.8%	80.2%
0.25	0.25	440	72.8%	27.2%
0.30	0.25	180	99.2%	0.8%
0.20	0.15	140	99.7%	0.3%
0.15	0.10	120	99.6%	0.4%

**Table 4:** Simulation Results

$\alpha = .05$ and $\beta = .05$				
$\theta_1(p_1, p_2)$		Median	Decision	
$p_1$	$p_2$	Sample Size	Unblind	Remain Blinded
0.15	0.25	1500	0.4%	99.6%
0.20	0.25	1500	6.7%	93.7%
0.25	0.25	1220	54.8%	45.2%
0.30	0.25	320	97.6%	2.4%
0.20	0.15	260	99.2%	0.8%
0.15	0.10	240	99.6%	0.6%

**Table 5:** Simulation Results

$\alpha = .1$ and $\beta = .05$				
$\theta_1(\rho_1, \rho_2)$		Median	Decision	
$\rho_1$	$\rho_2$	Sample Size	Unblind	Remain Blinded
0.15	0.25	1500	1.2%	98.8%
0.20	0.25	1500	12.2%	87.8%
0.25	0.25	760	64.1%	35.9%
0.30	0.25	240	98.4%	1.6%
0.20	0.15	200	99.5%	0.5%
0.15	0.10	180	99.8%	0.2%

**Table 6:** Simulation Results

$\alpha = .2$ and $\beta = .05$				
$\theta_1(\rho_1, \rho_2)$		Median	Decision	
$\rho_1$	$\rho_2$	Sample Size	Unblind	Remain Blinded
0.15	0.25	1500	4.0%	96.0%
0.20	0.25	1500	21.2%	78.8%
0.25	0.25	740	73.2%	26.8%
0.30	0.25	160	99.2%	0.8%
0.20	0.15	140	99.7%	0.3%
0.15	0.10	120	99.9%	0.1%

**Table 9:** Simulation Results

$\alpha = .2$ and $\beta = .1$				
$\theta_1(\rho_1, \rho_2)$		Median	Decision	
$\rho_1$	$\rho_2$	Sample Size	Unblind	Remain Blinded
0.15	0.25	1500	4.4%	95.6%
0.20	0.25	1500	21.3%	78.7%
0.25	0.25	380	74.3%	25.7%
0.30	0.25	160	99.2%	0.8%
0.20	0.15	140	99.7%	0.3%
0.15	0.10	120	99.9%	0.1%

What safety error rate is the DSMB willing to live with. For example from Table 1, if the DSMB chose  $\alpha=0.05$  and  $\beta=10^{-7}$  then we would indicate to the DSMB that they would likely “unblind” the trial early 6.3% of the time over theoretical repetitions of the study, and “unblind” the trial early 97.6% of the time if there was a true safety concern. Note that if the new treatment was more successful than anticipated the “unblinding rate” goes down to 0.5%. If the DSMB wanted a very stringent rule then they might go with  $\alpha=0.2$

**Table 7:** Simulation Results

$\alpha = .05$ and $\beta = .1$				
$\theta_1(\rho_1, \rho_2)$		Median	Decision	
$\rho_1$	$\rho_2$	Sample Size	Unblind	Remain Blinded
0.15	0.25	1500	0.4%	99.6%
0.20	0.25	1500	7.1%	92.9%
0.25	0.25	1120	56.3%	43.7%
0.30	0.25	320	97.8%	2.2%
0.20	0.15	260	99.0%	1.0%
0.15	0.10	240	99.6%	0.4%

**Table 8:** Simulation Results

$\alpha = .1$ and $\beta = .1$				
$\theta_1(\rho_1, \rho_2)$		Median	Decision	
$\rho_1$	$\rho_2$	Sample Size	Unblind	Remain Blinded
0.15	0.25	1500	1.6%	98.4%
0.20	0.25	1500	12.3%	87.8%
0.25	0.25	740	65.3%	34.7%
0.30	0.25	240	98.9%	1.1%
0.20	0.15	200	99.6%	0.4%
0.15	0.10	180	99.8%	0.2%

and  $\beta=10^{-7}$ , e.g. see Table 3. A possible tradeoff would be to choose  $\alpha=0.1$  and  $\beta=0.1$  in Table 8, where we would unblind the study at rate 12.3% if we were close to the proportions from which the study was planned around, and unblind the study 98.9% of the time if we were close to the safety hypothesis.

### DCA-MALA Trial Example

The following text appeared(modified slightly for this paper) in each DCA-MALA DSMB report following the adoption of the sequential test called the safety decision monitoring rule(SDMR) in the DCA-MALA trial. “After every 20 subjects (one randomization blocking unit) have completed the study the biostatistics coordinating center will “update” the SDMR and recommend to the DSMB to either remain blinded to the treatment assignment or to unblind the treatment assignment due to a significant increase in the mortality rate, beyond random noise, of the DCA group relative to

the placebo group. In addition to the SDMR, mortality data will always be monitored on case-by-case basis, i.e. if a sequential series of anomalous deaths occur in any given block the DSMB will be notified immediately, regardless of the SDMR”.

The original trial was designed to enroll  $n=1500$  subjects if accrual through two interim analysis reached 100%. Unfortunately, the trial was terminated early after only  $n=123$  subjects had completed the study due to problems stemming from failing to meet accrual milestones set by the sponsors. The low accrual rates were directly related to unusual dry spells occurring during the rainy seasons when malaria is prevalent. However, there was enough data gathered in order to demonstrate how the SDMR rule works in reality.

In order to illustrate the new SDMR to the DCA-MALA DSMB prior to their approval or disapproval, the following examples were presented to the committee members. The method was illustrated using various “made-up” outcomes, which were similar to what we anticipated might possibly occur during the DCA-MALA trial given

$$H_0 : \theta_0 = \text{efficacy} (p_1 = .19, p_2 = .25),$$

$$H_1 : \theta_1 = \text{safety} (p_1 = .28, p_2 = .25),$$

The results are presented in Tables 10 – 14. After every  $N_i = 10+10$  subjects completed the study  $\lambda_i$  was calculated along with the corresponding recommendation: “Remain Blinded” or “Unblind the Study.” The example given in Table 14 is the only case where DCA mortality was consistently lower than placebo mortality and hence the decision was always to “Remain Blinded.” In all other examples the decision to unblind the study was a function of the true underlying mortality rates. In our opinion these simple examples provided to the committee helped illustrate the utility of the SDMR and thus they ultimately endorsed its implementation.

## Results for the DCA-MALA Trial

In this section we illustrate how the SPRT decision rule worked within the context of the DCA-MALA trial through 120 subjects given  $\alpha=0.20$  and  $\beta=10^{-8}$  corresponding to boundaries of  $A=9.9999999$  and  $B=0.0000111$ . The trial was terminated due to financial circumstances after  $n=123$  subjects were enrolled. Hence, the final 3 subject’s data were not included in the safety monitoring statistic illustrated here.

Let  $\hat{p}_1$  and  $\hat{p}_2$  denote the estimated mortality rates in the DCA and placebo treatment groups for each block of 20 subjects. As can be seen the value of  $\lambda_i$  started to “drift” toward  $A=9.9999999$  as the imbalance in mortality rates favored DCA and then started to “drift” back towards  $B=0.0000111$  as the mortality rates became more balanced. The simplicity of programming this method is illustrated via the SAS program used to carry out the calculations given in Appendix B.

## Conclusion

In this note we presented a statistical decision rule for data safety monitoring purposes when the primary efficacy and primary safety endpoint of a clinical trial is mortality. This rule was designed for ease of interpretation by DSMB members with little or no formal statistical training. The goal of this method is to provide a means of controlling the approximate Type I error control for the efficacy analysis, while monitoring safety in a continuous fashion. As was discussed above, the method may be modified to accommodate more complex designs. Future work will involve studying the probability theory behind the utilization of different sequential bounds for efficacy and safety such this information can be incorporated into the sample size parameters during the design phase of the trial.

**Table 10:** Treatment trending towards being unsafe

Sample Size	DCA No. of deaths per 10	Placebo No. of deaths per 10	Decision
20	2	2	Remain Blinded
40	3	2	Remain Blinded
60	4	0	Unblind the study

**Table 12:** Treatment consistently worse.

Sample Size	DCA No. of deaths per 10	Placebo No. of deaths per 10	Decision
20	3	2	Remain Blinded
40	3	2	Remain Blinded
60	3	2	Remain Blinded
80	3	2	Remain Blinded
100	3	2	Remain Blinded
120	3	2	Unblind the Study

**Table 14:** Treatment consistently better

Sample Size	DCA No. of deaths per 10	Placebo No. of deaths per 10	Decision
20	1	2	Remain Blinded
40	1	2	Remain Blinded
60	1	2	Remain Blinded
80	1	2	Remain Blinded
100	1	2	Remain Blinded
120	1	2	Remain Blinded
140	1	2	Remain Blinded
160	1	2	Remain Blinded
180	1	2	Remain Blinded
200	1	2	Remain Blinded
220	1	2	Remain Blinded

**Acknowledgements**

Special thanks to David Harrington for suggesting developing many of the ideas included in this manuscript and for Sandy Zientek in the manuscript preparation. Dr. Hutson's work is partially supported by a NYSTAR Faculty Development Grant.

**Table 11:** Treatment consistently worse

Sample Size	DCA No. of deaths per 10	Placebo No. of deaths per 10	Decision
20	4	2	Remain Blinded
40	4	2	Remain Blinded
60	4	2	Remain Blinded
80	4	2	Unblind the study

**Table 13:** Spikes in the treatment death rate

Sample Size	DCA No. of deaths per 10	Placebo No. of deaths per 10	Decision
20	1	2	Remain Blinded
40	1	2	Remain Blinded
60	1	2	Remain Blinded
80	8	2	Remain Blinded
100	1	2	Remain Blinded
120	8	2	Unblind the study

**Table 15:** Results from the DCA-MALA Study

<i>i</i>	$\lambda_i$	Decision	$\hat{p}_1$	$\hat{p}_2$
1	1.40995	remain blinded	0.2	0.1
2	2.91669	remain blinded	0.1	0.0
3	3.89147	remain blinded	0.3	0.2
4	2.04235	remain blinded	0.0	0.1
5	2.12806	remain blinded	0.2	0.2
6	1.11686	remain blinded	0.0	0.1

**Appendix A: SPRT Statistical Details**

Let  $p_1$  and  $p_2$  denote the true underlying mortality rates in the DCA and placebo treatment groups, respectively. For our application we have developed a modified SPRT statistical decision rule which will test the following simple hypotheses after every 20 subjects have been enrolled, namely,

$$H_0 : \theta_0 = \arcsin \sqrt{p_1} - \arcsin \sqrt{p_2} ,$$



$$H_1 : \theta_1 = \arcsin \sqrt{p_1} - \arcsin \sqrt{p_2} .$$

The values for  $p_1$  and  $p_2$  were chosen based upon the original study design.

**Test Statistic.** Let  $\hat{p}_1 = (\# \text{ mortalities in the DCA arm}/n_{1i})$  and  $\hat{p}_2 = (\# \text{ mortalities in the placebo arm}/n_{2i})$ , where  $n_{1i}$  denotes the number of subjects per block  $i$  on DCA and  $n_{2i}$  denotes the number of subjects per block  $i$  on placebo,  $i=1, 2, \dots, 75$ . In most instances  $n_{1i}=n_{2i}=10$ . Even though the sample size per block is relatively small, the following asymptotic distributional results will serve the purpose for the SPRT decision rule,

$$\arcsin \sqrt{\hat{p}_{1i}} \sim AN \left( \arcsin \sqrt{p_1}, \frac{1}{4n_{1i}} \right), \tag{A.1}$$

$$\text{and } \arcsin \sqrt{\hat{p}_{2i}} \sim AN \left( \arcsin \sqrt{p_2}, \frac{1}{4n_{2i}} \right). \tag{A.2}$$

Next we define the SPRT test statistic

$$z_i = \frac{\arcsin \sqrt{\hat{p}_{1i}} - \arcsin \sqrt{\hat{p}_{2i}} - \theta}{\sqrt{\frac{1}{4n_{1i}} + \frac{1}{4n_{2i}}}}, \tag{A.3}$$

which will be monitored continuously in time as defined by block  $i$ . It is easily shown that  $z_i \sim AN(\theta_j, 1)$ , where

the distribution of  $z_i$  will be denoted  $f_{\theta}(z_i)$ , under  $H_j$ ,  $j=0,1$ , respectively. Based upon  $f_{\theta}(z_i)$  the SPRT test decision rule now follows straightforward from the classical methods. Calculate

$$\lambda_i = \frac{f_{\theta}(z_i)}{f_{\theta_0}(z_i)} \tag{A.4}$$

where  $f_{\theta}(z) = \prod_{i=1}^m f_{\theta}(z_i)$  and  $m = 1, 2, \dots$ , denotes the current block. If  $\lambda_i > A$  then stop and reject  $H_0$ . If  $\lambda_i < B$  then stop and accept  $H_0$ , else continue to monitor the study.

\renewcommand{\baselinestretch}{1}

**Appendix B: SDMR SAS Program**

```
proc format;
value dform 0='nblind'
           1='recalculate'
           2='remain blinded';
data simul;
input d1 d2 n1 n2;
n1=10;
n2=10;
cards;
2 1
1 0
3 2
0 1
2 2
0 1
;
data check;set simul;
retain f0 f1;
label p1='DCA death rate'
      p2='placebo death rate';
p1=d1/n1;
p2=d2/n2;

/*****Set Parameters*****/
hp1=.19; /*active under null*/
hp11=.28; /*active under alternative*/
hp2=.25; /*placebo*/
mu0=arsin(sqrt(hp1))-arsin(sqrt(hp2));
mu1=arsin(sqrt(hp11))-arsin(sqrt(hp2));
alpha=.1;
beta=.00001;
/*****

if _n_=1 then do;
f0=1;f1=1;
end;

/*calculate test statistic */
z0=(arsin(sqrt(p1))-arsin(sqrt(p2))-
mu0)/sqrt(1/(4*n1)+1/(4*n2));
z1=(arsin(sqrt(p11))-arsin(sqrt(p2))-
mu1)/sqrt(1/(4*n1)+1/(4*n2));
pi=2*acos(0);
f0=f0*exp(-z0*z0/2)/sqrt(2*pi*(1/(4*n1)+1/(4*n2)));
f1=f1*exp(-z1*z1/2)/sqrt(2*pi*(1/(4*n1)+1/(4*n2)));
lambda=f1/f0;
a=(1-beta)/alpha;b=beta/(1-alpha);
dec=2;
if lambda>a then dec=0;
if lambda<b then dec=1;
format dec dform.;

proc print;
var lambda a b dec p1 p2;
```

## References

1. Abraham Wald. *Sequential Analysis*. John Wiley & Sons Inc., New York, 1947.
2. Krishna, S., Nagaraja, R., Planche, T., Agbenyega, T., Bedo-Addo, G., Ansong, D., Owusa-Oforia, A., Shroads, A. L., Henderson, G., Hutson, A., Derendorf, H., Stacpoole, S. Population pharmacokinetics of intramuscular quinine in children with severe malaria. *J Clin Endocrinol Metabol* 2000; 85: 1569-76.
3. Expert Working Group (Efficacy) of the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). *Guideline for Industry, Structure and Content of Clinical Study Reports*, July 1996.
4. U.S. Department of Health and Human Services. Food and Drug Administration. International Conference on Harmonization; Guidance on Statistical Principles in Clinical Trials; Availability. *Federal Register* 1998; 63:49583-49598.