

# Sequential patient recruitment monitoring in multi-center clinical trials

Dong-Yun Kim<sup>1,a</sup>, Sung-Min Han<sup>b</sup>, Marston Youngblood Jr<sup>c</sup>

<sup>a</sup>National Heart, Lung and Blood Institute / National Institutes of Health, USA;

<sup>b</sup>Open Source Electronic Health Record Alliance (OSEHRA), USA;

<sup>c</sup>The University of North Carolina at Chapel Hill, USA

---

## Abstract

We propose Sequential Patient Recruitment Monitoring (SPRM), a new monitoring procedure for patient recruitment in a clinical trial. Based on the sequential probability ratio test using improved stopping boundaries by Woodroffe, the method allows for continuous monitoring of the rate of enrollment. It gives an early warning when the recruitment is unlikely to achieve the target enrollment. The packet data approach combined with the Central Limit Theorem makes the method robust to the distribution of the recruitment entry pattern. A straight-forward application of the counting process framework can be used to estimate the probability to achieve the target enrollment under the assumption that the current trend continues. The required extension of the recruitment period can also be derived for a given confidence level. SPRM is a new, continuous patient recruitment monitoring tool that provides an opportunity for corrective action in a timely manner. It is suitable for the modern, centralized data management environment and requires minimal effort to maintain. We illustrate this method using real data from two well-known, multicenter, phase III clinical trials.

**Keywords:** patient recruitment, continuous monitoring, clinical trial, sequential probability ratio test

---

## 1. Introduction

In a clinical trial, successful recruitment of patients is important. Since inadequate recruitment results in reduced sample size and may weaken the validity of the scientific findings, maintaining a steady stream of patient accrual is crucial. In fact, in a large clinical trial involving multiple sites, a periodic report of recruitment statistics is an integral part of clinical trial management.

At the design stage the project time and desired sample size are fixed to achieve the power of the statistical test associated with the primary endpoint. Throughout the trial, investigators monitor the enrollment process to see whether the number of subjects is projected to reach the goal. According to Lasagna's law (van der Wouden *et al.*, 2007), however, the researchers tend to be overly optimistic about the number of eligible patients in a given population. In Carlisle *et al.* (2015), out of 2,579 phase II and III clinical trials registered at the National Library of Medicine and ended in 2011, 19% were either terminated due to insufficient patient recruitment or ended with less than 85% of the planned enrollment. Scoggins and Ramsey (2010) report that more than 40% of the clinical trials sponsored by the National Cancer Institute did not meet the recruitment target. Rengerink (2014) estimates that

---

<sup>1</sup> Corresponding author: Mathematical Statistician, Office of Biostatistics Research, National Heart, Lung and Blood Institute, National Institutes of Health, 6701 Rockledge Drive, Bethesda, MD 20817, USA. E-mail: [dong-yun.kim@nih.gov](mailto:dong-yun.kim@nih.gov)

almost one half of the clinical trials in the Netherlands Trial Register with enrollment dates ending between 2005 and 2010, failed to meet the 80% of the target enrollment.

There are numerous reasons that the patient enrollment does not go as planned. No matter what causes the delay in the recruitment process, efficient monitoring tools give researchers a window of opportunity to detect and address the problem in a timely manner.

Currently there are a number of methods for patient recruitment monitoring. A simple but popular method uses ad-hoc rules that consist of a set of fixed proportions. At any given time during the recruitment period, recruitment is deemed “on the way” if the cumulative number of participating patients is at least 75% of the target number, prorated at that point in time. For example, suppose the recruitment plan is to randomize 400 patients in a 24-month period. Then the target enrollment at the six-month mark is 100 randomized patients. If the current patient count is at least 75, it is considered satisfactory. This heuristic approach reflects that there is a substantial variation in actual recruitment with the expectation that the recruitment will pick up pace later on.

In other approaches, prorated target numbers serve as the basis for recruitment indices. See Corregano *et al.* (2015) and Rojavin (2005). Deterministic approaches are easy to understand and implement. However, it is unclear how such guidelines are useful to assess the likelihood of reaching the target by the end of the recruitment period if the recruitment effort is found unsatisfactory.

There are recruitment monitoring tools based on probability models and stochastic processes. Carter (2004) used the Poisson distribution to model the distribution of the accrual period. Zhang and Lai (2011) used fractional Brown motion, while Zhang and Long (2010) used a non-homogeneous Poisson process and cubic B-spline to model the accrual distribution. Anisimov and Fedorov (2007) applied the Poisson-gamma model to multi-center clinical trials, and Gajewski *et al.* (2008) used a Bayesian approach. For a survey of various approaches, refer to Zhang and Long (2012). Also see Heitjan *et al.* (2015), Anisimov (2016).

These stochastic monitoring tools are a significant improvement over deterministic approaches, but they often share the same weakness: the stochastic models often require theoretical assumptions to be held during the entire recruitment process. In practice, these assumptions are difficult to verify prior to actual recruitment. Besides, according to Rengerink (2014), 11% of clinical trials changed the inclusion criteria to make it easier to meet the recruitment goal. Other protocol modifications that affect the sample size and enrollment period are also not uncommon, and it is not easy to incorporate such changes into existing deterministic or stochastic models.

In a different approach, Haidich and Ioannidis (2001) point out that the enrollment in the first month or two is strongly correlated with subsequent accrual. They argue that early enrollment records provide strong evidence for the feasibility of required accrual. However, such an assessment is subjective by nature and cannot be easily formulated to provide an actionable monitoring tool. In a large clinical trial, multiple sites start recruitment at different times and accrual rates so the initial assessment is more complicated than in a single center trial. Moreover, patient recruitment is often nonlinear, and feedback and interventions are quite common.

In this paper, we propose Sequential Patient Recruitment Monitoring (SPRM), a new monitoring method based on the sequential probability ratio test. The sequential test was originally developed during the Second World War for military applications (Wald, 1947). Compared to other statistical tests with the same type I and type II errors, it minimizes the expected sample size. SPRM inherits the desirable properties of this fully sequential approach, thus suitable for continuous monitoring. SPRM uses boundaries by Woodroffe (1982). Based on nonlinear renewal theory, these boundaries are an improvement over Wald boundaries.

By organizing recruitment entry times into a series of packet data, i.e., non-overlapping groups of

observations, SPRM requires only minimal assumptions on the accrual distribution. Consequently, it is applicable to a wide range of recruitment scenarios.

The monitoring method allows for timely feedback on whether enrollment is going as planned. If the enrollment rate falls significantly below a specified level during monitoring, it triggers a warning. At this point, SPRM can be used to estimate (1) the size of shortfall, (2) the probability of reaching the target enrollment, and (3) possible extension of recruitment period, based on the assumption that the recent trend that triggered the warning will continue. The monitoring method can also suggest a required enrollment rate for the remainder of the term to reach the target with high probability. This information could be useful for boosting enrollment by giving incentives to existing sites or adding new sites. At reset, hypotheses are updated to reflect these decisions and SPRM resumes monitoring, admitting a new set of patient recruitment data.

However, our approach does not call for a drastic change in clinical trial design or sample size in an adaptive fashion. In fact, SPRM assumes that the enrollment goal is fixed by the original design. The purpose of SPRM is more modest in that its primary function is to inform the users whether that goal is realistic, based on the evidence from data.

In the next section we formally introduce SPRM. In Section 3, we illustrate the method using actual recruitment data from two well-known, phase III clinical trials. In Section 4, we discuss the key features, limitations and recommendations.

## 2. Methods

### 2.1. Setup

Consider a clinical trial with target sample size  $n_0$  and recruitment period  $[0, T_0]$ . Suppose  $n_0$  and  $T_0$  are pre-determined before the recruitment starts. We wish to monitor the accrual process to examine whether  $n_0$  subjects are likely to be recruited by time  $T_0$ . Let  $\tau_i$  denote the time the  $i$ th patient enters the clinical trial and define  $\tau_0 = 0$ . Let  $Y_i = \tau_i - \tau_{i-1}$  be the inter-arrival time between the  $i$ th and  $(i-1)$ th patients,  $i = 1, 2, 3, \dots$ . Assume that  $Y_i$  is iid with some common distribution  $F$  with  $E(Y_i) = \mu$  and  $\text{Var}(Y_i) = \sigma^2$  where  $\mu$  and  $\sigma^2$  are positive finite parameters. We assume that  $\mu$  is unknown but  $\sigma^2$  is known. Suppose we wish to test

$$H_0 : \mu = \frac{T_0}{n_0} \quad \text{vs.} \quad H_1 : \mu = \frac{T_0}{n_0} \cdot \frac{1}{\delta}, \quad (2.1)$$

where  $\delta$  is a design parameter with  $0 < \delta < 1$ .

For  $j = 1, 2, \dots$ , let  $X_j = (1/m) \sum_{i=1}^m Y_{i+m(j-1)}$  denote the sample mean of the  $j$ th “packet” of size  $m$ . Let

$$Z_j = \frac{X_j - \mu}{\sigma / \sqrt{m}}$$

and  $\theta = E(Z_j)$ . By re-parameterization, the hypotheses (2.1) are equivalent to

$$H_0 : \theta = 0 \quad \text{vs.} \quad H_1 : \theta = \theta_1, \quad (2.2)$$

where

$$\theta_1 = \frac{\sqrt{m} T_0}{\sigma n_0} \left( \frac{1}{\delta} - 1 \right), \quad \theta_1 > 0. \quad (2.3)$$

Table 1: Woodroffe boundaries

$\theta$	$\gamma_0, \gamma_1$	$a, b$
0.2	0.89	2.88
0.4	0.79	2.76
0.6	0.71	2.65
0.8	0.63	2.53
1.0	0.56	2.42
1.2	0.50	2.30
1.4	0.45	2.19
1.6	0.40	2.08

Let  $G_0$  and  $G_1$  denote the distribution functions for  $N(0, 1)$  and  $N(\theta_1, 1)$ , respectively. By the Central Limit Theorem, the approximate distributions of  $Z_j$  under  $H_0$  and  $H_1$  are  $G_0$  and  $G_1$ . Let

$$g(z) = \frac{dG_1}{dG_0}(z)$$

Then the log-likelihood ratio test statistic  $\ell_n$  given  $z_1, z_2, \dots$  is

$$\ell_n := \sum_{j=1}^n \log g(z_j) = \sum_{j=1}^n \theta_1 \left( z_j - \frac{\theta_1}{2} \right). \quad (2.4)$$

Let  $R_1 = \inf \{n \geq 1 : \ell_n < -b \text{ or } \ell_n > a\}$  for some positive constants  $a$  and  $b$ .  $H_0$  is rejected if and only if  $\ell_{R_1} > a$ . Similarly,  $H_0$  is accepted if and only if  $\ell_{R_1} < -b$ . The boundaries  $a$  and  $b$  are calculated from Theorems 3.1 and 3.3, and Example 3.1 in Woodroffe (1982). Also see Siegmund (1975) and Lorden (1977).

Table 1 lists  $a$  and  $b$  for selected values of  $\theta$  when  $\alpha = \beta = 0.05$ . More generally,  $a = \log(\gamma_1/\alpha)$  and  $b = \log(\gamma_0/\beta)$  where  $\gamma_0$  and  $\gamma_1$  are the Laplace transformations of the excess limiting distribution under the null and alternative hypothesis, respectively, both evaluated at 1.

Under SPRM, after a decision is made by either accepting or rejecting the null hypothesis, a new test is proposed with updated hypotheses and a new parameter

$$\theta_2 = \frac{\sqrt{m}}{\sigma} \frac{T_1}{n_1} \left( \frac{1}{\delta} - 1 \right),$$

where  $T_1$  and  $n_1$  are remaining recruitment time and sample size, respectively. This assumes that  $\tau_{R_1} \leq T_0$  and  $R_1 \leq n_0/m$ . Now the SPRM will test  $H_0 : \theta = 0$  versus  $H_1 : \theta = \theta_2$  based on  $\ell_n$  with  $\theta_1$  replaced by  $\theta_2$  in (2.4). Note that  $\ell_n$  uses only the recruitment data newly available after the last decision. Subsequent hypotheses and tests are similarly defined.

## 2.2. Woodroffe boundaries for the normal distribution

Wald boundaries for the sequential test (Wald, 1947) have been around for a long time, and they are popular due to simplicity and ease of use. However, these stopping boundaries are asymptotically incorrect by a constant factor, as shown in Theorem 3.1, Woodroffe (1982). Consequently, Wald boundaries produce a wider interval than necessary. The test based on Wald boundaries tends to be somewhat conservative in that the empirical error probabilities are smaller than the nominal values. One of the practical implications is that it takes a longer wait until a decision is made, and it may miss the warning signal entirely. For comparison, Table 2 lists the sample mean and sample standard

Table 2: Optimal stopping time

$\theta$	$E_1(N_{Wr})$	$sd(N_{Wr})$	$E_1(N_{Wa})$	$sd(N_{Wa})$
0.2	137.51	99.63	144.21	103.54
0.4	33.83	24.08	37.26	26.30
0.6	15.28	10.65	17.52	11.90
0.8	8.71	6.09	10.34	6.99
1.0	5.74	3.97	7.06	4.67
1.2	4.02	2.71	5.14	3.31
1.4	3.06	2.02	3.99	2.52
1.6	2.42	1.58	3.23	2.00

deviation of the optimal stopping time for the two boundaries when the data were generated from  $N(\theta, 1)$ . Each value was calculated based on 10,000 simulations. In the table, the sample sizes  $N_{Wr}$  and  $N_{Wa}$  are associated with Woodroffe and Wald boundaries, respectively. Woodroffe boundaries are superior to Wald boundaries because the former has smaller mean and standard deviation than the latter for a wide range of  $\theta$ . It may appear that the gain in sample size is modest, but the unit is the number of packets. The difference of just two more packets requires 60 more observations when the packet size is 30, which may mean weeks or even months of additional waiting for the signal. Simulation studies (not included) also indicate that empirical type I and type II error probabilities of Woodroffe boundaries are closer to the nominal values of  $\alpha$  and  $\beta$ .

Woodroffe boundaries for the normal distribution assume the standard deviation is known. In practice, this is rarely the case so it should be estimated from data. Rather than starting monitoring from the very beginning, we recommend using a portion of the initial enrollment data for the parameter estimation before the formal monitoring starts. In a large, multi-center clinical trial, individual centers often start at different times and enrollment rates. So usually it takes awhile for the recruiting process to be stabilized, and allowing a lead-in time makes practical sense in this respect.

It is desirable to strike a balance between more precision for estimation and a shorter wait until monitoring starts. With lack of systematic approach to suggest an optimal starting time, we suggest the first 10 to 15% of the target enrollment to be used for initial estimation. This ad-hoc number is based on simulation studies and our field experience.

Let  $\hat{\sigma}$  denote the estimate of  $\sigma$ . Then the parameter  $\theta_1$  in (2.3) is understood as

$$\theta_1 = \frac{\sqrt{m} T_0}{\hat{\sigma} n_0} \left( \frac{1}{\delta} - 1 \right), \quad (2.5)$$

where  $T_0$  and  $n_0$  denote the remaining accrual period and the recruitment target minus the number of already enrolled patients. These notations will be used throughout the remainder of the paper.

### 2.3. Probability estimation and suggested rate guideline at reset

If the null hypothesis is rejected, it is an indication that the accrual process is significantly underperforming. When this happens, a quantity of interest is the probability of reaching the target enrollment by the end of the recruitment period, under the assumption that the current trend will hold for the remainder of the term.

To estimate this probability, first recall that  $T_1, n_1$  denote the remaining time and remaining number of new recruits needed, respectively. Let  $N(t)$  denote the number of new recruits within a time period of length  $t > 0$  and  $S_n$  denote the time until  $n$  patients are recruited. By the property of the

Table 3: Adjustment factor

$c \setminus n_1$	100	300	500	700	900	1100	1900
0.5	1.08	1.05	1.04	1.03	1.03	1.02	1.02
1.0	1.16	1.09	1.07	1.06	1.05	1.05	1.04
1.5	1.25	1.14	1.11	1.09	1.08	1.07	1.06
2.0	1.33	1.19	1.15	1.12	1.11	1.10	1.07

renewal process (Taylor and Karlin, 1998) and the Central Limit Theorem,

$$P(N(T_1) \geq n_1) \approx \Phi\left(\frac{T_1 - n_1\mu_1}{\sqrt{n_1}\sigma}\right), \quad (2.6)$$

where  $\mu_1 = T_0/(n_0\delta)$  is the parameter under the alternative hypothesis and  $\Phi(\cdot)$  is the distribution function for the standard normal distribution.

At the time of rejection, one may also be interested in the enrollment rate needed to recruit the remaining patients with probability  $p_0$ . Let  $q_0$  denote the  $100p_0^{\text{th}}$  percentile of the standard normal distribution. Assume that  $\sigma = c\mu$  for some positive constant  $c$ , which is called the coefficient of variation. Then the desired average waiting time between patient enrollment is

$$\mu = \frac{T_1}{n_1} \cdot \frac{1}{1 + q_0c/\sqrt{n_1}}. \quad (2.7)$$

Note that the first term on the right-hand side of equation (2.7) is the nominal value to be maintained for the remaining period. Due to the uncertainty inherent in the process, the average waiting time should be adjusted by the reciprocal of the adjustment factor  $A(c, n_1, p_0) = A = 1 + q_0c/\sqrt{n_1}$ .

In practice, it may be easier to think in terms of the average number of new recruits per day. In such cases, the desired average daily rate is  $\lambda^* = n_1A/T_1$ . Table 3 lists the values of the adjustment factor for selected combinations of  $c$  and  $n_1$  for  $p_0 = 0.95$ . This table gives a practical guideline for the daily recruitment numbers if the higher assurance (95%) of reaching the target is essential for the clinical trial.

#### 2.4. Expected shortfall and extension of recruitment period

After the null hypothesis is rejected, the accrual rate for the remaining period should be increased to reach the target recruitment goal. This may be difficult to accomplish due to several reasons, e.g., rarity of the disease or stringent inclusion/exclusion criteria. Then an extension of the recruitment period could be an option.

Suppose after the first rejection, one wishes to determine how much extension is needed to reach the target recruitment  $n_0$  with some high probability  $p_0$ . From equation (2.7), substitute the lefthand side by  $\mu_1 = T_0/(n_0\delta)$  and solve for  $T_1$ . After simple algebra, the additional recruitment time (ART) is

$$\text{ART} = \left(\frac{n_1A}{n_0\delta} - 1\right)T_0 + \tau_{R_1}. \quad (2.8)$$

The performance of equation (2.8) was evaluated based on 10,000 computer simulations. For this, we considered a typical scenario that one may encounter in a large multi-center clinical trial. We assumed one year for the recruitment period. Patient recruitment process was simulated by a Poisson process with intensity specified in the alternative hypothesis in (2.1) using design parameter  $\delta = 0.75$ .

Table 4: Additional recruitment time (days)

$n_0$	ART	sd(ART)	$\hat{n}_0$	sd( $\hat{n}_0$ )	$\hat{\tau}_{R_1}$	Power
1,000	133.8	5.6	1025	30.2	60.1	0.93
2,000	130.5	2.8	2036	43.4	30.0	0.92
3,000	129.0	1.9	3046	53.8	20.0	0.92
4,000	128.0	1.4	4052	63.0	15.0	0.92

The coefficient of variation was set to be 1.0 for simplicity. The Woodroffe stopping boundaries were determined using  $\alpha = 0.05$  and  $\beta = 0.1$ . Table 4 summarizes the results. In the table,  $\hat{\tau}_{R_1}$  denotes the average time at rejection, and the required minimum extension was estimated by equation (2.8), which would ensure to reach the target accrual 80% of the time. The estimated sample size at stopping was about 123, and  $\hat{n}_0$  denotes the average number of eventual recruitments when the suggested extension was employed. Computer simulations indicate that equation (2.8) gives a simple guideline when the extension of the recruitment period is considered during a clinical trial.

Another way to increase recruitment is to add more recruitment sites without extending the accrual time. Suppose the accrual rate remains approximately the same for the remaining duration. In equation (2.7),  $\mu^*$  is the desired average waiting time between patients so the reciprocal  $1/\mu^*$  is the average number of recruitment per day, assuming the day is the time unit. If the null hypothesis is rejected, the current average number of recruitment per day is  $n_0\delta/T_0$ . Hence  $n_0\delta T_1/T_0$  is the expected number of total recruitment by the end of remaining recruitment period  $T_0$ . The expected shortfall (ES) is given by the following formula.

$$ES = n_1 - \frac{n_0\delta T_1}{T_0}. \quad (2.9)$$

We can use equation (2.9) when additional sites are required to achieve the target recruitment. If either the extension of accrual period or expected shortfall is contemplated at the time of later rejection, formulas (2.8) and (2.9) should be modified accordingly in a recursive fashion.

### 3. Illustration

In this section we illustrate the sequential monitoring method using enrollment data from two phase III, multi-center clinical trials. For both cases, we used the first 15% of the entry data for the initial setup and estimation. We also used 5% Type I and II errors,  $\delta = 0.75$  and the packet size of 30.

#### 3.1. ENhancing recovery in coronary heart disease trial

ENhancing recovery in coronary heart disease (ENRICHHD) (Berkman *et al.*, 2003) was a two-arm, randomized, multi-center clinical trial that studied the effect of cognitive behavior therapy on survival and quality of life for those patients who already had a myocardial infarction (MI). The composite primary endpoint was nonfatal MI or all-cause mortality. The original enrollment plan was to recruit 3,000 patients in three years; the randomization started in November 1996 and ended in October 1999. The actual enrollment was 2,481 patients or 82.7% of the original plan. SPRM gives the first warning signal at day 418, less than four months after monitoring started on day 301. The second warning followed the first by three months (day 504). Additionally SPRM issued warning signals five times afterward. In Figure 1, solid triangles indicate the times of warning signals during the recruitment period. The broken vertical line marks the start of monitoring.

At the second rejection on day 504, one may wish to estimate the probability of reaching the target enrollment at the end of the period. To do so, we use the formula (2.6). With  $T_1 = 591$  days

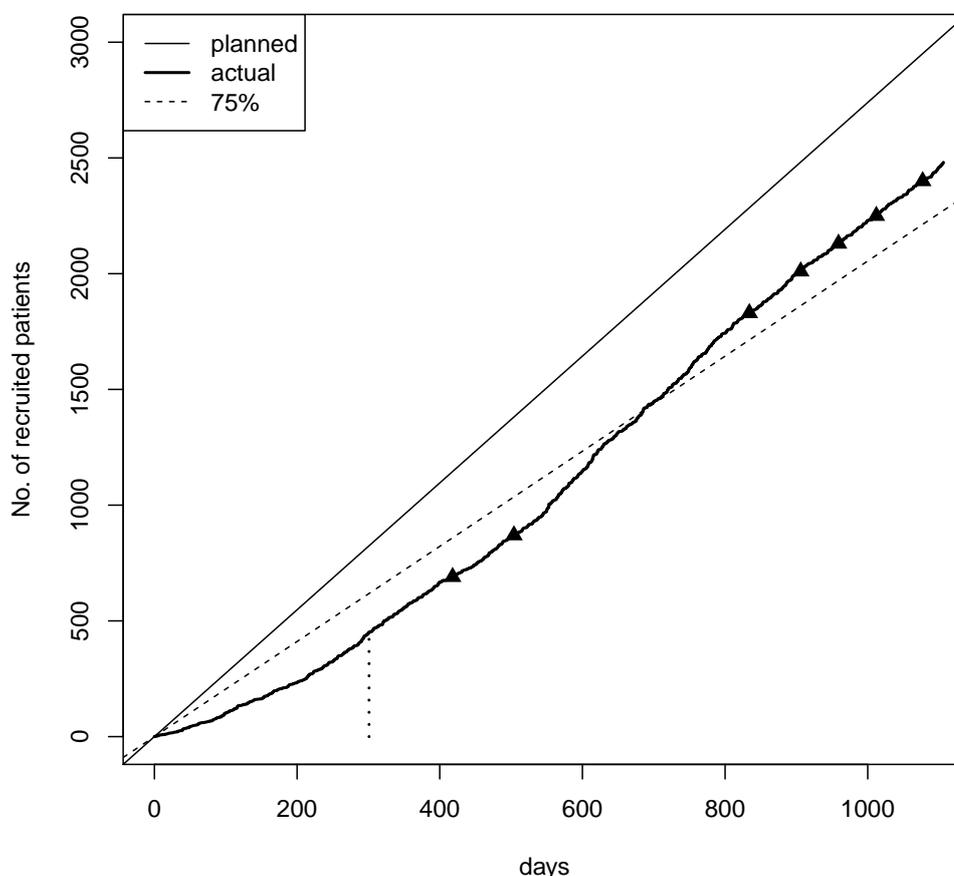


Figure 1: *SPRM Plot: ENhancing Recovery In Coronary Heart Disease*. Solid triangles indicate the times of warning signal. The broken vertical line marks the start of monitoring. *SPRM = Sequential Patient Recruitment Monitoring*.

left to additionally recruit  $n_1 = 2,130$  patients, the probability is  $\Phi(-7.77) \approx 0$  when  $\mu_1 = 0.39$ , the parameter under the alternative hypothesis is used. The expected shortfall calculated from (2.9) is 615 patients while the actual number is 519 at the end of the accrual period.

Suppose an extension of accrual period is considered at that point. Utilizing formula (2.8), the accrual period should be extended approximately nine months beyond the original three-year period to recruit the target 3,000 patients with 80% of chance. This is in line with the actual accrual rate of 827 patients per year of the ENRICHD trial.

### 3.2. Beta-blocker Heart Attack trial

Beta-blocker Heart Attack Trial (BHAT) (Byington, 1984) was a double-blind, randomized, placebo-controlled multi-center clinical trial where 30 U.S. sites and 1 Canadian site participated. The purpose of the study was to investigate whether long-term propranolol therapy significantly reduces all-cause mortality for patients with acute MI. The original enrollment plan for the trial was to recruit 4,200

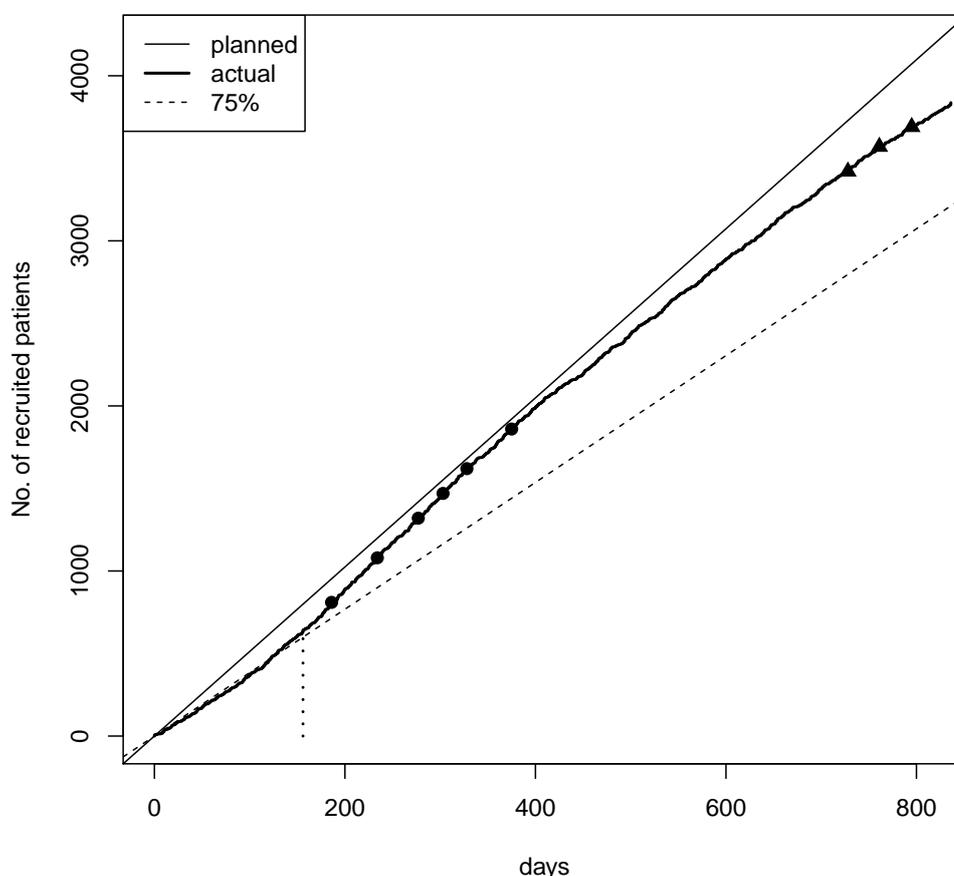


Figure 2: SPRM Plot: Beta-blocker Heart Attack Trial. The broken vertical line marks the start of monitoring. Solid circles indicate the times of signal that the recruitment is on target. Solid triangles indicate the times of warning signal. SPRM = Sequential Patient Recruitment Monitoring.

patients in 27 months. Randomization started in June 1978 and ended in October 1980. The actual number of enrolled patients was 3,837 or 91.4% of the original plan.

In this illustration, the formal monitoring commenced five months after recruitment started. In Figure 2, between the sixth month and one year mark, SPRM signaled six times (marked by solid circles) to indicate the recruitment was on target to meet the 4,200 patient goal, if the trend continued. Starting at the two year mark, however, SPRM gave three warning signals in succession (marked by solid triangles). These are a strong indication that the recruitment would likely fall short of the target enrollment.

The estimated probability of reaching the target enrollment, at the time of the last acceptance (day 375) is 0.508. Such a seemingly low probability follows the fact that the nominal target rate concerns the average inter-arrival time of successive recruitment. In a clinical trial where the success of reaching the target is crucial, one may desire a higher probability, say 0.95 instead. Since the nominal target rate is 53 new recruitment per 10 days and the adjustment factor is estimated 1.07, the

desired daily rate is 56 patients to ensure success with a 95% chance. The desired target rate at the specified probability could serve as a practical guideline during the recruitment period.

#### 4. Discussion

Statistical patient monitoring methods often assume a certain form for the underlying distribution of the time between successive recruitments. In practice, this is difficult to know ahead or even during the trial because the recruitment patterns rely on individual clinical trials, characteristic of the disease, severity, inclusion/exclusion criteria among others. Besides, a misspecification of the underlying distribution often leads to poor inference.

One way to solve the problem is to use the sequence of sample means calculated from the series of non-overlapping packets of enrollment data. Suppose the observations are iid with a common distribution with finite mean and variance. For an adequate packet size  $m$ , the Central Limit Theorem allows us to treat packet sample means as if they were iid observations from a normal distribution. Simulation study suggests that the sequential monitoring works well with even modest  $m$ , say between 5 to 30.

An obvious downside of this approach is that it reduces the effective sample size by the factor of  $m$ . Consequently, to implement SPRM, the target enrollment need to be fairly large, say at least 200 observations.

In the illustrations, we chose 0.75 for the design parameter  $\delta$ . This is because 75% of the target is often used as an informal benchmark in data coordinating center (DCC) reports of many clinical trials that we have observed. Although other choices of  $\delta$  are certainly possible, we think 0.75 serves well for a general monitoring purpose. If actual accrual is far too low, say 50% of the goal is being achieved, it would be evident from the DCC reports.

There are other limitations in the current implementation. We recommended an ad-hoc range of 10–15% of planned recruitment for the timing of the initial estimation of the standard deviation, a nuisance parameter. A more formal procedure would be desirable to determine the optimal start time for monitoring. In the current approach, we did not consider patient dropouts. This is because dropouts are expected in the course of a clinical trial so it is usually taken into account when the initial recruitment goal is determined. However, unexpected excessive dropouts have a negative impact on recruitment so it may well be incorporated into the monitoring process.

To estimate the shortfall at the end of the accrual period, we used the parameter under alternative hypothesis when the test rejected the null hypothesis. Another possible approach is to use the estimate for the average daily accrual rate using data collected up to that point. In such a case, it may require bias correction because the estimation is made after the sequential test. This is a well-known fact in sequential inference and there are standard bias correction methods available. We did not take the approach here, but it would be an interesting comparison to our method.

The repeated significance test may have a theoretical edge over our current approach when the parameters are unknown. However, one could also argue that the goal of SPRM is to give a simple but practical guideline for stakeholders in clinical trials.

In spite of these limitations, SPRM is a useful addition to the toolbox that is available to the sponsor, study investigators and DCC. SPRM can help stakeholders make an informed and timely decision regarding patient recruitment in a clinical trial.

## Acknowledgements

The authors wish to thank Drs. Nancy Geller and Eric Leifer at the Office of Biostatistics Research, NHLBI for reading the early draft and providing insightful comments.

The views expressed in this article are those of the authors and do not represent the views of NHLBI, NIH, OSEHRA, or the University of North Carolina.

## References

- Anisimov VV (2016). Discussion on the paper "Real-time prediction of clinical trial enrollment and event counts: a review" by DF Heitjan, Z Ge, and GS Ying, *Contemporary Clinical Trials*, **46**, 7–10.
- Anisimov VV and Fedorov VV (2007). Modelling, prediction, and adaptive adjustment of recruitment in multicentre trials, *Statistics in Medicine*, **26**, 4958–4975.
- Berkman LF, Blumenthal J, Burg M, *et al.* (2003). Effects of treating depression and low perceived social support on clinical events after myocardial infarction: the Enhancing Recovery in Coronary Heart Disease Patients (ENRICH) Randomized Trial, *Journal of the American Medical Association*, **289**, 3106–3116.
- Beta-Blocker Heart Attack Trial (BHAT) <https://biolincc.nhlbi.nih.gov/studies/bhat/?q=bhat> (accessed 17 January 2017).
- Byington RP (1984). Beta-blocker heart attack trial: design, method, and baseline results, *Controlled Clinical Trial*, **5**, 382–437.
- Carlisle B, Kimmelman J, Ramsay T, and MacKinnon N (2015). Unsuccessful trial accrual and human subjects protections: an empirical analysis of recently closed trials, *Clinical Trials*, **12**, 77–83.
- Carter RE (2004). Application of stochastic processes to participant recruitment in clinical trials, *Controlled Clinical Trials*, **25**, 429–436.
- Corregano L, Bastert K, Correa da Rosa J, and Kost RG (2015). Accrual index: a real - time measure of the timeliness of clinical study enrollment, *Clinical and Translational Science*, **8**, 655–661.
- ENRICH protocol version 7.0. <https://biolincc.nhlbi.nih.gov/static/studies/enrichd/Protocol.pdf> (accessed 16 January 2017).
- Gajewski BJ, Simon SD, and Carlson SE (2008). Predicting accrual in clinical trials with Bayesian posterior predictive distributions, *Statistics in Medicine*, **27**, 2328–2340.
- Haidich A and Ioannidis JP (2001). Patterns of patient enrollment in randomized controlled trials, *Journal of Clinical Epidemiology*, **54**, 877–883.
- Heitjan DF, Ge Z, and Ying GS (2015). Real-time prediction of clinical trial enrollment and event counts: a review, *Contemporary Clinical Trials*, **45**, 26–33.
- Lorden G (1977). Nearly-optimal sequential tests for finitely many parameter values, *The Annals of Statistics*, **5**, 1–21.
- Rengerink KO (2014). Embedding trials in evidence-based clinical practice (PhD thesis), University of Amsterdam, Amsterdam.
- Rojavin MA (2005). Recruitment index as a measure of patient recruitment activity in clinical trials, *Contemporary Clinical Trials*, **26**, 552–556.
- Scoggins JF and Ramsey SD (2010). A national cancer clinical trial system for the 21st century: reinvigorating the NCI cooperative group program, *Journal of National Cancer Institute*, **102**, 1371.
- Siegmund D (1975). Error probabilities and average sample number of the sequential probability ratio test, *Journal of the Royal Statistical Society. Series B (Methodological)*, **37**, 394–401.

- Taylor HM and Karlin S (1998). *An Introduction to Stochastic Modeling* (3rd ed), Academic Press, San Diego.
- van der Wouden JC, Blankenstein AH, Huibers MJ, van der Windt DA, Stalman WA, and Verhagen AP (2007). Survey among 78 studies showed that Lasagna's law holds in Dutch primary care research, *Journal of Clinical Epidemiology*, **60**, 819–824.
- Wald A (1947). *Sequential Analysis*, John Wiley, New York.
- Woodroffe M (1982). Non linear renewal theory in sequential analysis, *CBMS-NSF Regional Conference Series in Applied Mathematics*, Society for Industrial and Applied Mathematics.
- Zhang Q and Lai D (2011). Fractional Brownian motion and long term clinical trial recruitment, *Journal of Statistical Planning and Inference*, **141**, 1783–1788.
- Zhang X and Long Q (2010). Stochastic modeling and prediction for accrual in clinical trials, *Statistics in Medicine*, **29**, 649–658.
- Zhang X and Long Q (2012). Modeling and prediction of subject accrual and event times in clinical trials: a systematic review, *Clinical Trials*, **9**, 681–688.

Received March 2, 2018; Revised August 17, 2018; Accepted August 17, 2018