PREDICTION OF ACCRUAL CLOSURE DATE IN MULTI-CENTER CLINICAL TRIALS WITH POISSON PROCESS MODELS

by

Yuan Kong

B.S., Peking Union Medical College, China, 1999

M.S., in Biostatistics, University of Pittsburgh, 2003

Submitted to the Graduate Faculty of

the Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2009

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This dissertation was presented

by

Yuan Kong

It was defended on

July 31st 2009

and approved by

Gong Tang, PhD, Assistant Professor, Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh

Joseph P. Costantino, Dr.PH., Professor, Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh

Chung-Chou Ho Chang, PhD, Research Associate Professor, Department of Medicine, School of Medicine, University of Pittsburgh

Lan Kong, PhD, Assistant Professor, Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh

Dissertation Director: Gong Tang, PhD, Assistant Professor, Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh

PREDICTION OF ACCRUAL CLOSURE DATE IN MULTI-CENTER CLINICAL TRIALS WITH POISSON PROCESS MODELS

Yuan Kong, PhD

University of Pittsburgh, 2009

Objective: To develop a systematic statistical approach to estimate accrual closure date in large scale multi-center clinical trials or large public health studies. It is relevant to the research in public health.

Background: In a typical multi-center cancer clinical trial or large public health study, sample size is predetermined to achieve desired power and study participants are enrolled from hundreds of satellite sites. As the accrual is closing to the target size, the coordinating data center needs to project an accrual closure date based on observed accrual pattern and notify participating sites several weeks in advance. In the past, projections were simply based on some crude assessment and conservative measures were incorporated in order to achieve the target accrual size. The resulted excessive accrual size usually leads to unnecessary budget increase considering that the coordinating center needs to pay thousands of dollars for each accrued participant.

Method: For multi-center clinical trials, there is very small probability for a site to accrue a patient during a short period and mostly the accrual from different sites is mostly independent from each other. Therefore, the overall accrual could be modeled by a Poisson process. Based on accrual data collected up to a time point, a Poisson process-based method was used to analyze the past accrual pattern. Combining with assumption on the future accruing pattern, two methods were proposed here to predict the accrual closure date. The estimates and their confidence intervals were used to guide clinical practice. . The proposed methods were illustrated through analysis of accrual data from NSABP trials B-38 and C-08.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

First of all, I would like to thank my advisor, Dr. Gong Tang, for his guidance, encouragement and support throughout years of my graduate study. I have learned not only sound ways of doing research from him, but also being an honest and caring person. I truly am grateful to Dr. Tang's patient revision of my proposal and this dissertation. I would also like to thank my committee members Dr. Joseph P. Costantino, Dr. Chung-Chou Ho Chang and Dr. Lan Kong for their valuable suggestions and thoughtful comments throughout the course of this dissertation.

It has been my pleasure to be able to use the data from NSABP biostatistics center. I would like to express my appreciation to NSABP.

Many Thanks to my parents for their years of care and encouragement and to my friends for their continued support. Finally, I wish to express special thanks and to dedicate this work to my husband, Haobo, for his endless care. I simply could not have taken the initial step of my study nor have gone this far without him. I would like to share my happiness with him. I also want to dedicate this work to my son, Nathan, who was born on May 9th this year.

1.0 INTRODUCTION

A clinical trial is a prospective study that compares the effects and other aspects of one or more treatments or interventions against a control in human beings. It has become an integral component of development of new drugs and interventions in the pharmaceutical industry and health science studies. Study participants are usually enrolled and followed up for a pre-determined period. The effects of designated interventions on the health of participants are then recorded and studied. In general, there are four types of clinical trials. Phase I clinical trials are used to determine the maximal tolerated dose. Phase II trials are used to evaluate the effects and side effects of the new interventions. Phase III trials are designed to compare effects and safety of the new interventions against a control. Phase IV trials sometimes are conducted to follow up new interventions that are shown effective in phase III trials. Among them, phase III trials are usually the most visible ones because they are large and costly, and their results will determine the fate of new interventions that have been shown promising in earlier stages.

The spending on clinical trials is enormous and ever increasing, although patient care costs for clinical trials are not appreciably higher than costs for patients not enrolled in trials. Compared to phase I and II trials, phase III trials are relatively more expensive. The average cost of a phase III clinical trial could potentially exceed $20,000 per patient. There could be many reasons for the higher spending on the clinical trials. Some are necessary costs which are hard to avoid, such as costs on patient registration, treatment, subsequent health management, and data collection. However, some costs, such as cost due to extra enrollment beyond the target sample size, may be lessened. In a multi-center phase III trial, the sample size is pre-determined in order to achieve sufficient power for its primary hypothesis. Participants are often recruited from hundreds of study centers such as hospitals,

community clinics and cancer institutes over time. Because of the complexity of recruiting process, it is almost impossible for the coordinating center, especially of cancer clinical trials, to determine a cut-off date on accrual and eventually reach the target sample size exactly. It has been the usual practice to choose a cut-off date so that reaching the target sample size can be "guaranteed" though this often leads to extra and sometimes excessive accruals. Extra accruals to a certain degree are helpful because they would fill in for participants who later on withdraw consent or lose to follow up. However, excessively extra accrual would increase unnecessary financial burden on the coordinating centers and their sponsors.

Based on some trials we observed at National Surgical Adjuvant Breast and Bowel Project (NSABP), one of the Cooperative Groups of National Cancer Institute, additional patients were accrued beyond the target sample sizes. The primary interest of this paper is to provide a systematic and statistical method to determine the closure date so that we could reach the target sample size at the end of the closure day without recruiting too many additional patients although sometimes accruing several additional patients will be helpful to guarantee the enough sample size. There could be ineligible patients who can not pass the screening, especially in cancer clinical trials. Therefore having a few more patients would help to maintain sufficient power. But investigators usually wish to limit the additional accrual to a certain extend.

The Clinical Trials Cooperative Groups of the National Cancer Institute consist of researchers at affiliated institutions to jointly design and carry out multi-center cancer treatment and prevention clinical trials. These seventeen groups accrue approximately 20,000 new patients at an annual basis. The majority of those clinical trials are phase III trials, and often thousands of patients are required in a single trial. Usually, the sample sizes are determined prior to the initiation of trials in order to achieve sufficient power for the primary hypotheses. The NSABP study B-38 was designed as a phase III clinical trial to compare the treatment efficacy of three adjuvant chemotherapy regimens for node-positive breast cancer patients. A target sample size of 4,800 was required to achieve sufficient power for the primary hypothesis. Another NSABP randomized phase III trial, C-08, was to compare two adjuvant regimens on colon cancer patients, 2632 was the target sample size.

In practice, patients are accrued from hundreds of participating institutions or sites over time. As the accrual is approaching the target sample size, the coordinating data center is required to project an accrual closure date, based on the observed accrual pattern after the first patient entering the trial, and notify participating members of this closure date several weeks in advance. The reason to make the prediction several weeks in advance is because patient accrual is a complicated process, in which the participating centers need days or even weeks to evaluate the eligibility of potential participants and acquire their consent to the study. After the coordinate center finds out the target sample size is reached, it cannot request the participating centers to discontinue patient enrollment abruptly because some potential participants have already entered the accrual process though not completed. Such estimated closure dates are usually based on a simple assessment of the observed accrual pattern, such as averaging accrual for the past few months. Then the length of further accrual period is determined by the required amount of further accrual to reach the target sample size and the estimated accrual rate. Conservative measures, such as adding a few more days of accrual, are also incorporated in order to reach the target accrual and guarantee sufficient power for the primary analysis. In B-38, projection of the accrual closure date occurred about 6 weeks before the closure date and 4,894 patients were actually accrued at the end of the closure date. In C-08, projection of the accrual closure date also occurred about 6 weeks before the closure date and 2710 patients were actually accrued at the end of the closure date. Some of the extra accruals in the B-38 study and C-08 study are helpful because dozens of participants withdrew consent and another few were deemed ineligible afterwards.

In many occasions, such overflow of accrual in large multi-institutional clinical trials usually leads to unnecessary budget increase, considering that the cost for treatment and management on an average cancer patient takes tens of thousands of dollars. If there are about 2.5% more patients than necessary, each year this will add about tens of million of dollars in cost solely for managing patients who participate trials of these Cooperative Groups. In order to obtain a more precise prediction of the closure date, we pursue a systematic and statistical approach to predict closure date in a multi-center clinical trial. At first a Poisson process model was used to model the observed accrual pattern for a specific

trial up to the time when the coordinating data center is required to project an accrual closure date. Then functions, such as mean and quantiles, of a future date when the accrual reaches the target sample size would be estimated based on obtained inference results on the observed accrual pattern and assumptions on the future accrual pattern. Related issues such as model diagnostics and sensitivity analyses, under alternative assumptions of the future accrual pattern, were investigated here. Extension to under more complex settings would be considered. The proposed method would supply a simple, rigorous and robust tool for efficient accrual in multi-center clinical trials.

1.1 THE MOTIVATING EXAMPLE: NSABP B-38 TRIAL

In 2004, a phase III trial B-38 was designed by NSABP to compare the treatment efficacy of three adjuvant chemotherapy regimens for node-positive breast cancer patients. The accrual size of 4,800 was pre-determined in order to achieve sufficient power for the primary hypothesis that the investigative arm, dose-dense doxorubicin and cyclophosphamide (AC) followed by paclitaxel plus gemcitabine, improved the disease-free survival over the other two arms, docetaxel plus AC, and dose-dense AC followed by paclitaxel alone, respectively. Patients were accrued from hundreds of participating institutions or sites over time. The protocol was opened on October 1, 2004, and the first patient entered this trial on November 3, 2004. As the cumulative accrual was closing to the target sample size, a total of 4465 patients were accrued by March 20, 2007. The coordinating data center was required to project an accrual closure date for this trial based on the accrual pattern before March 20th, 2007 and send it to participating sites several weeks in advance. At that time, May 3 of 2007 was determined as the closure date. This prediction was based on the average daily accrual during the 3-month period before March 20, 2007 and three more days were added in order to make sure that the target sample size would be reached. The speculation of future accrual rate was crude. Eventually, 4894 patients were randomized in this trial at the end and the final accrual was 94 over the target sample size.

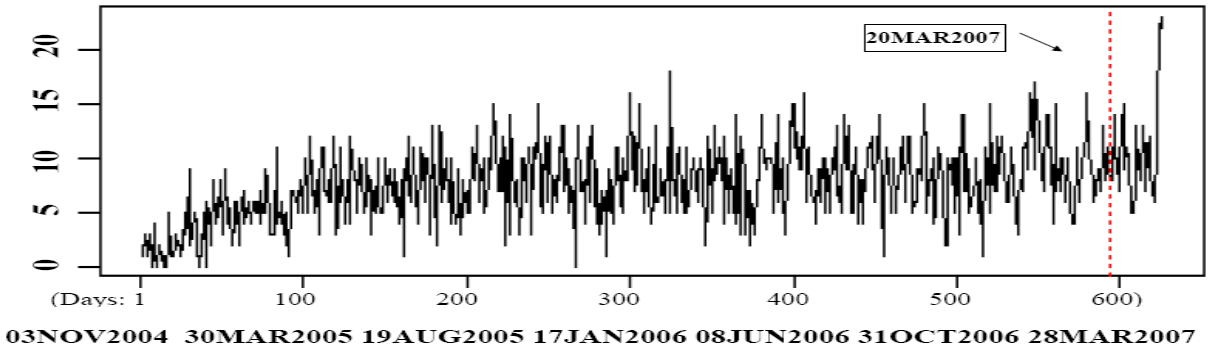

Figure 1: Overall Daily Accrual For NSABP Trial B-38.

Figure 1 shows the daily accrual in B-38 over time. Several interesting phenomena were observed. First, during the first 4 or 5 months, the accrual rate was increasing steadily, reflecting that it took time for the participating institutions to fulfill regulatory requirements and start recruiting patients for this trial. Secondly, the accrual pattern became relatively stable after March of 2005. At the last, during the last few days of the accrual period prior to the designated closure date, the daily accrual increased to about twice of the average daily accrual during the middle of the accrual period. Information from this study supplied an excellent opportunity for us to study the accrual patterns for typical cancer clinical trials and prediction of the trial closure dates.

2.0 EXISTING APPROACHES ON THE PREDICTION OF TRIAL CLOSURE

In a multi-center clinical trial, a target sample size is pre-determined in order to achieve reasonable power for the primary hypothesis. Whether the target accrual size can be reached in a timely fashion often has great impact on the time frame of the definitive analysis. A fast accrual will lead to an early closure date and subsequent early dissemination of the study results to the public. Usually at the design stage of a trial, an accrual pattern is anticipated based on accrual information from past trials on similar study population. Then the investigators have an idea about how long the accrual will take and when the definitive analysis will be expected. In clinical practice, the first patient may not be entered until days or weeks after the initiation of the trial because it takes time for the participating institutions to prepare themselves for regulatory requirements and put patients on the trial. With more and more institutions gear up for the trial, the accrual rate starts to increase steadily until reaching a stable level. If this stable accrual rate is higher than the anticipated rate, the accrual can be completed earlier than expected and the primary analysis can be carried out earlier. Otherwise, the primary analysis may be delayed because of slower accrual. In cases when the actual accrual rate is much slower than anticipated, interventions such as soliciting aggressive efforts from participating centers are necessary. Sometimes a trial has to be closed when possible interventions to speed up accrual have been exhausted but the accrual rate is still far below the anticipated level.

If the accrual rate is satisfactory and the cumulative accrual is approaching the target sample size, the coordinating center usually needs to predict a date when the accrual will reach the target sample size and notify the participating institutions in advance. The member institutions can keep recruiting patients until the end of this disseminated closure date. If

the eventual accrual size is a lot less than the target sample size, then the study may be considerably under-powered than planned. If the eventual accrual size is much larger than the target sample size, then unnecessary expenditure on accrual incurs. Therefore it is important for the coordinating center to appropriately determine the closure date when the target sample size will be reached without incurring unnecessarily extra accruals.

Current practice in the prediction of the closure date is mostly carried out as following: first calculate the number of accruals that is needed to reach the target sample size, then divide this number by the average daily or weekly accrual rate during past few months. This method will supply a crude estimator of time period for further accrual with assuming that the future accrual will follow the same rate as observed in the past few months. Sometimes a few extra days or weeks are added as a conservative and precautionary measure. This approach is simple and reasonable but is often rough and inadequate. During the past decade, Senn (1998) [15], Anisimov and Fedorow (2007) [3] and Gajewski et al. (2008) [9] provided some systematic procedures to predict the closure date based on Poisson process models for accrual patterns.

2.1 A POISSON PROCESS-BASED FRAMEWORK

In a typical multi-center phase III trial carried out by NSABP, hundreds of participating institutions recruited hundreds or thousands of patients during 2 to 4 years. Consequently, the chance for a single institution to recruit one patient during a short period such as a day would be small, let alone the chance to recruit more than one patient. Therefore, the total accrual from those institutions during a short period approximately follows a Poisson distribution (Taylor & Karlin, 1998 [16]). In general the patient accruals during disjoint time intervals are independent from each other. These properties lead to an observation that the accrual process approximately follows a Poisson process (Taylor & Karlin, 1998 [16]).

Consider a multi-institutional trial, the daily accrual from all participating institutions is denoted by $\{X(t), t = 1, 2, \ldots\}$. Assume that patient accrual follows a Poisson process with accrual rates $\{\lambda(t), t = 1, 2, \ldots\}$, where t indicates the time from initiation of the trial

or some specific date such as the date when the first patient is recruited. The cumulative accrual up to time t would be:

$$N(t) = \sum_{s=1}^{t} X(s).$$

Senn (1998) considered modeling the accrual process as a homogeneous Poisson process with constant rate λ. Denote $W_{N(t)+1}$ as the sojourn time from the arrival of the $N(t)$th patient to the next accrual. Then this sojourn time follows an exponential distribution with mean $1/\lambda$ and its density is

$$f(w) = \lambda e^{-\lambda w}.$$

For the prediction of accrual time, assume that N is the target sample size and λ is known, then the required time to recruit $N - N(t)$ after time $T = t$, or the waiting time till the target sample size, follows a gamma distribution with scale $1/\lambda$ and shape $N - N(t)$ (Taylor & Karlin, 1998 [16]; Senn, 1998 [15]).

Compared to some common methods for predicting the closure date in practice, such as those based on averaging accrual for the past few months, the method which Senn (1998) [15] proposed was based on a simple and proper statistical model for general practice. Although the model is rather simple without considering some complicated scenarios, this framework would supply a straightforward and useful starting point before more extensive exploration.

In practice, the accrual rate, λ, is often unknown and can be estimated based on prior accrual data $\{X(s), s = 1, 2, \ldots, t.\}$. However, Senn (1998) [15] did not give a clear guidance on how to estimate λ. It could be estimated based on some historical clinical trials which share some similarity at patient characteristics, treatment regimen and etc. The estimation could base on partial accrual pattern from the same clinical trial if we only need to predict the closure date after certain amount of patients accrued. Moreover, Senn (1998) [15] did not consider the possibility of the fluctuation of the accrual rate over time, like what we observed in the motivation example - NSABP trial B-38. The accrual rate increased at the first 4-5 months, and then it became relatively stable for a while until the last 5 days

when the accrual rate nearly doubled. This whole accrual process showed more variability than a homogeneous Poisson process. Furthermore, Senn (1998) [15] did not discuss how the variability of the estimated λ would affect the estimated mean time required for further accrual.

2.2 A BAYESIAN METHOD UNDER THE POISSON PROCESS-BASED FRAMEWORK

Gajewski et al. (2008) [9] proposed a Bayesian method to predict accrual closure date under Senn (1998) [15]'s Poisson process framework. Denotes t_i as the times when each new study participant enters the trial and assume that the study starts at time $t_0 = 0$. Denote $w_i = t_i - t_{i-1}$ as the waiting times from accruing the $(i-1)$th participant to accruing the ith participant. They assumed that the underlying Poisson process is homogeneous and consequently the waiting time, w_i, follows an exponential distribution with mean $1/\lambda$. Gajewski et al. (2008) [9] believed that such an assumption conformed to what they observed in clinical practice and also most clinical trials have prior information on the parameter $1/\lambda$.

In their Bayesian approach, the inverse gamma distribution was selected as the prior for λ with parameters (α, β). It is the conjugate prior. Gajewski et al. (2008) [9] proposed to determine the values of parameters for the prior distribution, or hyper-parameters, based on historical clinical data. If the prior parameters were identified based on some historical clinical trials from a similar population, a possible weight for the prior parameters was also discussed given the sample size of the historical clinical trials. Basically, if the historical trial has less target sample size than the current trial, then the prior can be weighted equal to the size of the previous sample size. Otherwise, the prior will be weighted less so it will not overwhelm the posterior distribution with data that are not directly related to the current trial.

Given the accrual data up to time t, $\mathcal{D}(t) = \{X(s), s = 1, 2, \ldots, t.\}$, the interest is to estimate the additional waiting time for accruing the remaining $N - N(t)$ patients. If the waiting time w_i follows an exponential distribution, the predictive distribution of an

unobserved $W_{N(t)+1}$ has a closed form. For more general cases, Gajewski et al. (2008) [9] derived its posterior distribution $p(\lambda \mid \mathcal{D}(t))$. In order to predict the length of further accrual time for reaching the target sample size, Gajewski et al. (2008) [9] proposed the following procedure to generate a random sample of time needed to reach the target sample size N:

(i) Draw λ from its posterior distribution $p(\lambda \mid \mathcal{D}(t))$.

(ii) Generate the waiting times $W_{N(t)+1}, W_{N(t)+2}, \ldots$, and W_N from exponential distributions with the drawn λ as parameter.

(iii) Sum up the simulated waiting times.

With repeating this procedure for numerous times, a distribution of the length of further accrual time could be presented for decision making. As a matter of fact, this procedure can be simplified by simulating the time to reaching the target sample size directly by a gamma distribution instead of a sum of numerous waiting times. It is because that essentially they assumed that the accrual followed a homogeneous Poisson process. Based on the theory of Poisson process, the required time to recruit the remaining patients $N - N(t)$, or the waiting time $W_{N(t)+1}$, follows a gamma distribution.

Gajewski et al. (2008) [9] also compared three different priors, such as only the information at the beginning of the study based on investigators' opinion, non-informative prior without any background information about the accrual pattern, and informative prior based on partially observed data. Based on the examples presented in the paper, Gajewski et al. (2008) [9] concluded that the greatest degree of uncertainty occurs with the non-informative prior.

Compare to some prediction of accrual rates made on ad hoc basis, this approach by Gajewski et al. (2008) [9] is a quantitative and rigorous. However, they also assumed that the accrual followed a homogeneous Poisson process. As we discussed, most of the trials show more variability than a homogeneous Poisson process. Although this procedure is straightforward and easy to implement, it over-simplified the accrual process. Also, the inference is a little complex and ambiguous in terms of computation and choice of the prior distribution of λ. Like all other models based on Bayesian method, selecting the appropriate prior distribution is sensitive and can be very subjective. In this paper, identifying the parameters for

the prior distribution is also very subjective since it is based on the investigator's personal knowledge and experience. Gajewski et al. (2008) [9] argued that a research would have at least some ideas about the projected accrual rate at the designing stage and their model also account for some uncertainty of the accrual rate since they asked questions about their confidence about the overall accrual process when preparing for the prior parameters, instead of their opinions about the accrual rates.

2.3 A POISSON-GAMMA MODEL

Anisimov and Fedorov (2007) [3] considered a more complex Poisson-gamma model. They argued that in real trials, the accrual rates varied across centers and the starting times for centers were different. The resulted accrual process would be more complicated than a homogeneous Poisson process. They proposed to model this variability by assuming that the accrual rates from participating centers follow a gamma distribution and patient arrival within each center follows a homogeneous Poisson process. It was called a Poisson-gamma model by Anisimov and Fedorov (2007) [3], or a compound Poisson process in the literature.

Denote $n_i(t)$ as the number of accruals at a participating institution or center i up to time t and $n(t) = \sum_{i=1}^{m} n_i(t)$ as the total number of patients accrued up to time t from all m centers. Furthermore, it was assumed that patients arrival at each center i followed a homogeneous Poisson process with a site-specific rate λ_i. These rates were unknown and assumed to follow a gamma distribution, $Ga(\alpha, \beta)$. Therefore the overall accrual up to time t is a Poisson process with the following rate:

$$\Lambda(t) = \sum_{i=1}^{m} \lambda_i 1\{t \geq u_i\},$$

where u_i is the time when accrual at institution i was initiated; $1\{t \geq u_i\} = 1$ when $t \geq u_i$, $1\{t \geq u_i\} = 0$ otherwise.

The overall accrual would be a heterogeneous Poisson process with instantaneous rate $\Lambda(t)$ which is a random variable and each newly initiated center would add an additional rate λ_i after u_i. However, u_i, the time when accrual at institution i is initiated is rarely

available. If assume that all centers are initiated simultaneously at the same time t_0 and the institution-specific rate λ_is are known, the overall accrual would become a homogeneous Poisson process with constant rate $\Lambda(t) \equiv \Lambda = \sum_{i=1}^{m} \lambda_i$.

Given observed accrual data up to time t, $\{x_i(s), i = 1, 2, \ldots, m;\ s = 1, 2, \ldots, t\}$, Anisimov and Fedorov (2007) [3] considered the prediction of the remaining accrual time under the assumption that all institutions would continue recruit without interruption and no more institutions would join after time t. Hence the future accrual would follow a homogeneous Poisson process with rate $\Lambda = \sum_{i=1}^{m} \lambda_i$. Based on the theory of homogeneous Poisson processes, the remaining recruitment time to reach the target sample size N follows a gamma distribution, $Ga(N - n(t), 1/\Lambda)$ when Λ is known.

With assuming that $\{\lambda_i\}$s are random and follow a gamma distribution $Ga(\alpha, \beta)$, the the future accrual process then follows a Poisson process with a random rate $\Lambda \sim Ga(m\alpha, \beta)$. After estimating the hyper-parameters using either maximum likelihood or method of moments, the distribution of Λ is known and so are the distribution of needed further accrual time, which follows $Ga(N - n(t), 1/\Lambda)$.

Anisimov and Fedorov (2007) [3] focused on the occupancy problem in multi-center clinical trials that not all participating centers contribute patients during any given period. Much of their discussion on estimation and prediction of future closure date centered on that concept.

The model by Anisimov and Fedorov (2007) [3] assume patients arrive at different centers according to Poisson process and the rates are a sample from a gamma distribution. They believe that their model reflects the natural variation in recruitment rates observed in practice and it was applied in several of completed trials. It can serve as a basic recruitment model. However, it can not model the complicated scenarios in the process of patient accrual like the high accrual rate at the end of the trial or the fluctuation of the accrual pattern over time. The method of predicting of remaining recruitment time is rather complex and is hard to apply in real life.

3.0 A FLEXIBLE POISSON PROCESS MODEL FOR PATIENT ACCRUAL IN MULTI-CENTER TRIALS

3.1 INTRODUCTION

In a typical large multi-center phase III cancer clinical trial, usually thousands of patients are required. Patients are recruited from hundreds of participating sites from various medical institutions or hospitals. During a work day, there is a small chance for a single institution to put a patient on the trial. Patient accrual during a certain time period is mostly independent from the accrual during another disjoint time period. The accruals from different sites are usually independent from each other as well. Therefore Poisson processes supply a natural tool to model the patient accrual for a phase III trial (Kingman, 1993 [13]; Taylor & Karlin, 1998 [16]; Senn, 1998 [15]).

After a multi-center trial is initiated, it usually takes several weeks for institutions to prepare themselves before the start of recruiting patients. They need to get IRB approval first before accrual and some sites may need to hire staff to operate the trial. Such gap is also differential among institutions, for example, teaching hospitals that are affiliated with research universities may have better trained doctors and staff members and subsequently shorter gap for gear-up than the community-based medical clinics. During this initiation period, the daily accrual rates increase slowly until all sites are ready to accrue patients. Therefore the homogeneous Poisson process model (Senn, 1998 [15]; Gajewski et al., 2008 [9]) cannot be applied under this circumstance. There is more variability than a homogeneous Poisson process, which assumes the accrual rates are constant from the beginning of the trial to the closure day.

Furthermore, the time of initiation for each participating institutions may not be recorded at the data coordinating center and this will make the application of the approach by Anisimov and Fedorov (2007) [3] rather difficult. They considered a homogenous Poisson process with a random accrual rates for each participate center, as all centers did not initiate simultaneously. Often times, the time when each institution is ready to accrue might not be available in the database since it is usually less important in the process of patient accrual.

After a closure date is predicted, often we can observe a dramatic increment in daily accrual during the last few days before the designated closure date. This phenomenon was observed in many NSABP trials, including B-38 and C-08, and it reflected that the participating institutions had tried their best to put their patients on the trial before the deadline because they believed that their patients would be benefited. All existing methods are not able to handle prediction for trials with this aspect.

In practice, when a closure date is predicted, the participating institutions are allowed to recruit by the end of that day and the data coordinating center cannot put a stop solely based on the cumulative accrual. Gajewski et al.(2008) [9], Anisimov and Fedorov (2007) [3] both tried to supply an estimator of the mean remaining accrual time or an empirical distribution of the remaining accrual time, provided that the exact arrival times for existing participants are known. However, in clinical practice, the arrival dates, rather than the exact arrival times, for existing participants are recorded. The coordinating centers also provide a predicted date, rather than a time point, as the trial closure date. Although it is a minor problem and those methods may supply a valid estimator by using the ceiling of the original estimator, its operating characteristic is not clear. Moreover, all of the exciting methods did not discuss on the model diagnostics of the required assumptions such as Poisson process and homogeneity of the accrual rate.

In this chapter, we present a flexible Poisson process-based model for describing patient accrual, two methods for prediction of the trial closure date, model diagnostics and some sensitivity analyses. These methods are illustrated through analysis of the accrual data from the NSABP B-38 and C-08 trials in Chapter 4.

3.2 A FLEXIBLE POISSON PROCESS MODEL FOR TRIAL ACCRUAL

We propose a flexible Poisson process to model the observed pattern of patient accrual. Consider a multi-institutional trial, the daily accrual of the coordinating center denotes by $\{X(t), t = 1, 2, \ldots\}$. Assume that the patient's arrival follows a Poisson process and the daily accrual rates are $\{\lambda(t), t = 1, 2, \ldots\}$, where t indicates the accrual days.

Assume that accrual needs a period of time, say t_0, to start and running up till reaching full potential at a later time t_1. Then the accrual becomes stable for a substantial period. When the cumulative accrual is close to the target size N, say at time $t_2 > t_1$, the coordinating center needs to assess the observed accrual pattern up to t_2 and predict a future date as the trial closure date in advance. It has also been widely observed that during the final few days prior to the designated closure date, the accrual rate is much higher than the past accrual rates. The following Poisson process model can be used for estimating the daily accrual rate $\lambda(t)$:

$$\lambda(t) = \begin{cases} 0, & \text{when } t \leq t_0; \\ \lambda(t;\beta), & \text{when } t_0 < t \leq t_1; \\ \mu, & \text{when } t_1 < t \leq t_3 - \Delta; \\ c\mu, & \text{when } t > t_3 - \Delta \end{cases} \tag{3.1}$$

where t_3 is the predicted closure date and Δ is the length of the short period with high accrual rate before t_3. If the initiation of participating institutions follows a homogeneous Poisson process before t_1, then the intensity rate $\lambda(t, \beta)$ during this period is a linear function of t. In practice, time point t_0 usually refers to the day before the first accrual; t_1 can be determined as the date when all of the participating sites have IRB approval and start recruiting patients for the trial, or be speculated as a date when the trial looks like having reached a stable accrual rate; t_2 is always known. t_2 is the time point when the coordinate center needs to predict the future closure date. The constant $c > 1$ could be speculated from accrual patterns of past trials that are similar to the current trial in terms of patient

characteristics and treatment regimen. However, it is also possible that the historical clinical trials are not available. Hence one way to evaluate the different values of c is to conduct sensitivity analysis. The accrual rate during $[t_0, t_1]$, $\lambda(t, \beta)$ is usually unknown and often assumed to be a linear function of t at the stage of trial design and planning.

Denote T, unknown and random, as the day when the accrual will surpass the target sample size N. In the following context, we would consider the prediction of T with modeling the accrual process as a nonhomogenous Poisson process as 3.1, the determination of t_1 when the accrual rates become constant, the model diagnostics on the assumptions of Poisson process, and the extension to more complex models.

3.3 INFERENCE BASED ON PAST ACCRUAL PATTERN

Suppose after the accrual started, at time t_2, where $t_2 > t_1$, we need to predict the future accrual closure date. First, we need to estimate μ. We could use the following likelihood based on the accrual from t_0 to t_2 to estimate μ. Consider the accrual pattern $\mathcal{D}(t) = \{X(t), t = 1, 2, \ldots, t_2\}$ from t_0 to t_2 has been observed.

If the accrual rate during $[t_0, t_1]$, $\lambda(t, \beta)$, is a linear function of t, then the likelihood function is:

$$\begin{aligned} L(\mu) &= \prod_{t=1}^{t_2} p(X(t) = x(t); \mu) \\ &= \prod_{t=t_0+1}^{t_2} p(X(t) = x(t); \mu, \delta) \\ &\propto \{ \prod_{t=t_0+1}^{t_1} exp(-\frac{t-t_0}{t_1-t_0}\mu)(\frac{t-t_0}{t_1-t_0}\mu)^{x(t)}\} \{ \prod_{t=t_1+1}^{t_2} exp(-\mu)\mu^{x(t)}\}. \end{aligned} \tag{3.2}$$

The likelihood has two components. The first part contributes from when the accrual increases linearly $[t_0 + 1, t_1]$ and the second part contributes from the constant accrual rate period, $[t_1 + 1, t_2]$.

Then, a log-likelihood function of $L(\mu)$ can be written in the form:

$$\begin{aligned} logL(\mu) &= \sum_{t=t_0+1}^{t_1} [-\frac{t-t_0}{t_1-t_0}\mu + x(t)\{\log\frac{t-t_0}{t_1-t_0} + \log\mu\}] + \sum_{t=t_1+1}^{t_2}\{-\mu + x(t)\log\mu\} \\ &\propto -\{\sum_{t=t_0+1}^{t_1}\frac{t-t_0}{t_1-t_0} + (t_2-t_1)\}\mu + \sum_{t=t_0+1}^{t_2} x(t)\log\mu \end{aligned}$$

Therefore the maximal likelihood estimator (MLE) of μ is:

$$\widehat{\mu} = \frac{\sum_{t=t_0+1}^{t_2} x(t)}{\sum_{t=t_0+1}^{t_1}\frac{t-t_0}{t_1-t_0} + (t_2-t_1)}$$

However, the assumption of a linear trend in accrual rates during the ramp-up period $[t_0+1, t_1]$ may not be appropriate. Then one may estimate μ based on the accrual data during the period $[t_1+1, t_2]$. The corresponding likelihood function of μ becomes:

$$\begin{aligned} L_a(\mu) &\propto \prod_{t=t_1+1}^{t_2} exp(-\mu)\mu^{x(t)} \qquad (3.3) \\ logL_a(\mu) &\propto \sum_{t=t_1+1}^{t_2}\{-\mu + x(t)\log\mu\} = -(t_2-t_1)\mu + \sum_{t_1+1}^{t_2} x(t)\log\mu. \end{aligned}$$

The corresponding MLE of μ is

$$\widehat{\mu}_a = \frac{\sum_{t=t_1+1}^{t_2} x(t)}{t_2-t_1}.$$

Also, we can get the observed Fisher information of μ as:

$$I_a(\mu) = \frac{1}{\mu^2}\sum_{t=t1+1}^{t_2} x(t)$$

Based on the Fisher information of μ, we can calculate the standard error of the ML estimator $\widehat{\mu}_a$ as

$$\text{se}(\widehat{\mu}_a) = I_a(\widehat{\mu}_a)^{-\frac{1}{2}} = \frac{\{\sum_{t=t_1+1}^{t_2} x(t)\}^{\frac{1}{2}}}{t_2-t_1}.$$

In practice, if the accrual rate stays stable after a certain time point, it would be simpler to estimate the constant accrual rate $\lambda(t) = \mu$ according to the likelihood $L_a(\mu)$. However, the choice of t_1 would affect the estimate and a procedure for finding a sensible choice of t_1 is important. How to select t_1 will be discussed later.

3.4 ESTIMATION OF c FROM A PAST TRIAL

As we observed from Figure 1, at the last few days before the designated closure date t_3, the participating sites usually try to get their patients in before the closure date because they believe their patients would be benefited from the trial. Denote Δ as the high accrual period. The accrual rate usually runs up during the final few days, where $t \in [t_3 - \Delta + 1, t_3]$. In Function 3.1, the accrual rates during the last few days is donated as $\lambda(t) = c\mu$, $c > 1$. Based on the similar completed accrual data from a past multicenter trial carried out by the same coordinating center, an estimate of c can be obtained from maximizing the following likelihood function:

$$\begin{aligned} L_c(\mu, c) &= \prod_{t=t_1+1}^{t_3-\Delta} exp(-\mu)\mu^{x(t)} \prod_{t=t_3-\Delta+1}^{t_3} exp(-c\mu)(c\mu)^{x(t)} \qquad (3.4) \\ logL_c(\mu, c) &= \sum_{t=t_1+1}^{t_3-\Delta} [-\mu + x(t)log\mu] - c\Delta\mu + \sum_{t=t_3-\Delta+1}^{t_3} x(t)(logc + log\mu) \\ &\propto -(t_3 - \Delta - t_1 + c\Delta)\mu + \sum_{t=t_3-\Delta+1}^{t_3} x(t)logc + \sum_{t_1+1}^{t_3} x(t)log\mu. \end{aligned}$$

Then the MLE of c is:

$$\widehat{c} = \frac{(t_3 - \Delta - t_1)\sum_{t=t_3-\Delta+1}^{T} x(t)}{\Delta \sum_{t=t_1+1}^{t_3-\Delta} x(t)}.$$

For estimating c, the historical trial needs to share some similar characteristics as the current running trial, such as disease status, treatment regimen, coordinate center, participants institutions and etc. It may also involve clinical knowledge on patient accrual in practice. However, it is always possible that such a similar completed clinical trial is not available. We could conduct some sensitivity analysis to determine how the predicted accrual closure date varies with potential choices of c. First, we could set c equal some potential values. With a selected t_1, we could project the corresponding closure dates and predicted total final accruals by those closure dates. Based on the variation in closure dates and total final accrual, the influence of c can be evaluated and then the value of c will be determined by the primary investigators. The sensitivity analysis will be illustrated in Chapter 4.

3.5 PROJECTION OF THE CLOSURE DATE

Suppose at time $t_2 > t_1$, about $N(t_2)$ patients have been randomized. We are interested in predicting the future closure date after calculating the estimated MLE of the accrual rate $\lambda(t) = \mu$ based on the observed accrual up to time t_2. The coordinating center needs to set a trial closure date T in order to notify participating sites of this closure date several weeks in advance. Assume that T is the date when the accrual surpasses the target accrual N, i.e., $T = \min\{t : N(t) \geq N\}$. Two methods are proposed here to determine the future closure date T. First, investigators are often interested in the expectation of T, $E(T)$, which is the average of time when the overall accrual would pass the target sample size N. Another value of interest is the quantiles of T, a future date $T_\alpha = min\{t : pr[T \leq t] \geq 1 - \alpha\}$. We could pick a small α, say $\alpha = 0.1$ or 0.05, so that at the closure date, we would be pretty sure that the sufficient sample size would be achieved.

3.5.1 Estimation of $E[T]$

The following algorithm is for calculating the expectation of the closure date, which is the average of time when the overall accrual would pass the target sample size.

Let $W = T - t_2$ denote as the additional accrual time after t_2 to reach the target sample size, then:

$$\begin{aligned}
E[W] &= \sum_{k=0}^{\infty} pr[W > k] \\
&= \sum_{k=0}^{\infty} pr[T - t_2 > k] = pr[T - t_2 > 0] + \sum_{k=1}^{\infty} pr[T - t_2 > k] \\
&= 1 + \sum_{k=1}^{\infty} pr[N(t_2 + k) < N] \\
&= 1 + \sum_{k=1}^{\infty} pr[N(t_2 + k) - N(t_2) < N - N(t_2)] \\
&= 1 + \sum_{k=1}^{\infty} \sum_{j=0}^{N-N(t_2)-1} pr[N(t_2 + k) - N(t_2) = j]
\end{aligned}$$

$$= 1 + \sum_{k=1}^{\infty} \sum_{j=0}^{N-N(t_2)-1} \frac{e^{-\mu_k}\mu_k^j}{j!}, \tag{3.5}$$

where

$$\mu_k = E[N(t_2+k) - N(t_2)] = \begin{cases} (k-\Delta)\mu + c\Delta\mu, & \text{if } k > \Delta; \\ ck\mu, & \text{otherwise.} \end{cases}$$

and Δ is the high accrual period, a few days before the trial closure date. The estimated expectation of T would be $\widehat{E(T)} = \widehat{E(W)} + t_2$.

The probability of the closure date is greater than t_2 or $T - t_2 > 0$ is 1. It is because that the total accrual $N(t_2)$ at t_2, must be less than the target sample size, so we need to accrue more patients and predict a closure date, which in turn means t_2 is before the closure date. Otherwise, if $N(t_2) = N$, then $t_2 = T$, it would be too late to inform the participating centers to stop accruing since the whole process of close a trial could take several weeks.

We could also calculate the standard error of the estimated expectation of T, $\widehat{E[T]}$, using Delta method. It equals the standard error of $\widehat{E(W)}$. We could calculate standard error of $\widehat{E(W)}$ first. Denote $\theta = E[W] = h(\mu)$, a natural estimator of $E[W]$ is $\widehat{\theta} = h(\widehat{\mu})$. Since $h(\cdot)$ is a smooth function, the standard error of $\widehat{\theta}$ can be estimated by the Delta method as the multiple of the standard error of $\widehat{\mu}$ and the absolute value of the derivation of $\widehat{\theta}$, $s.e.(\widehat{\mu}) \mid h'(\widehat{\mu}) \mid$. From the above representation of $\theta = h(\mu)$, the functional form of $h'(\mu)$ is:

$$\begin{aligned} h'(\mu) &= \sum_{k=1}^{\infty}[-e^{-\mu_k}\frac{d\mu_k}{d\mu} + \sum_{j=1}^{N-N(t_2)-1}\{je^{-\mu_k}\mu_k^{j-1} - e^{-\mu_k}\mu_k^j\}\frac{d\mu_k}{d\mu}/j!] \\ &= \sum_{k=1}^{\infty}\frac{d\mu_k}{d\mu}[\sum_{j=0}^{N-N(t_2)-2} e^{-\mu_k}\mu_k^j/j! - \sum_{j=0}^{N-N(t_2)-1} e^{-\mu_k}\mu_k^j/j!] \\ &= -\sum_{k=1}^{\infty}\frac{d\mu_k}{d\mu}\frac{e^{-\mu_k}\mu_k^{N-N(t_2)-1}}{(N-N(t_2)-1)!} \\ &= -\sum_{k=1}^{\Delta}\frac{cke^{-ck\mu}(ck\mu)^{N-N(t_2)-1}}{(N-N(t_2)-1)!} \\ &- \sum_{k=\Delta+1}^{\infty}\frac{(k-\Delta+c\Delta)e^{-\mu_k}\mu_k^{N-N(t_2)-1}}{(N-N(t_2)-1)!} \end{aligned} \tag{3.6}$$

The standard error of $\widehat{T}$ or $\widehat{W}$ can be calculated accordingly.

3.5.2 Estimation of T_α

Given a small value of α, the quantile of future closure time T_α can be estimated using the following algorithm. First, we calculate the CDF of T as a function of $\mu = \hat{\mu}$. Then, we could find out the smallest t to make $Pr(T \leq t)|_{\mu=\hat{\mu}}$ greater than or equal to $1-\alpha$. It could be integrated as: we have $(1-\alpha)\%$ confidence that the target sample size would be achieved by the end of the day T_α.

For any given $t > t_2$, let

$$\begin{aligned} 1-\alpha &\leq Pr(T \leq t) = pr[N(t) \geq N] \\ &= pr[N(t) - N(t_2) \geq N - N(t_2)] = 1 - pr[N(t) - N(t_2) < N - N(t_2)] \\ &= 1 - \sum_{k=0}^{N-N(t_2)-1} pr[N(t) - N(t_2) = k]. \end{aligned} \tag{3.7}$$

Note that, $N(t)$ the total accrual at the end of the trial closure date can be derived as the summation of patients accrued before t_2 which is known, patients accrued after $t_2 + 1$ but before and during the high accrual period Δ.

$$N(t) = N(t_2) + \sum_{s=t_2+1}^{t-\Delta} x(s) + \sum_{s=t-\Delta+1}^{t} x(s)$$

and the additional accrual from time t_2+1 to the closure time t follows a Poisson distribution with the parameter as $((t-\Delta-t_2)\mu + \Delta c\mu)$.

$$\sum_{s=t_2+1}^{t-\Delta} x(s) + \sum_{s=T-\Delta+1}^{T} x(s) \sim P((t-\Delta-t_2)\mu + \Delta c\mu)). \tag{3.8}$$

Therefore, the probability distribution function of the future accrual after t_2 is

$$\begin{aligned} pr[N(t) - N(t_2) = k] &= pr[\sum_{s=t_2+1}^{t-\Delta} x(s) + \sum_{s=t-\Delta+1}^{T} x(s) = k] \\ &= \frac{exp\{-(t-\Delta-t_2)\mu - c\Delta\mu\}\{(t-\Delta-t_2)\mu + c\Delta\mu\}^k}{k!}, \end{aligned}$$

where $k = 0, 1, \ldots, N - N(t_2) - 1$. Then substitute above function into Function 3.7, we could calculate the smallest T_α, where $T_\alpha = min\{t : pr[T \leq t] \geq 1-\alpha\}$ when μ is known.

With μ estimated based on observed accrual data $\mathcal{D}(t_2)$ up to t_2, the estimator $\widehat{\mu}$ can be used in the above formula to estimate T_α but its variability needs to be taken into account. A simple method for estimating the standard error of the estimator of T_α is the bootstrap method. The algorithm is presented as following:

(1) For the dth bootstrap sample, randomly sample $\{N(t_1+1), N(t_1+2), \ldots, N(t_2)\}$ with replacement and denote the re-sampling accrual data as $Y^{(d)}$.

(2) For the dth bootstrap sample, obtain $T_\alpha^{(d)}$. Repeat for $d = 1, 2, ..., D$, where D denotes a large number.

(3) Let $\bar{\widehat{T}}_\alpha = \frac{1}{D}\sum_{i=1}^{D} T_\alpha^{(d)}$. Then, the standard error of the T_α is: s.e.$(\widehat{T}_\alpha) = \sqrt{\frac{1}{D-1}\sum_{d=1}^{B}(T_\alpha^{(d)} - \bar{\widehat{T}}_\alpha)^2}$.

The confidence interval of the estimator of T_α can be calculated accordingly.

3.6 PREDICTION OF TOTAL NUMBER OF ACCRUAL GIVEN A CLOSURE DATE t_3

After the closure date t_3 is determined, either by $\widehat{E[T]}$ or $\widehat{T}_\alpha$, it would be also informative to present the expected total accrual by t_3. Because the future accrual $N(t_3) - N(t_2)$ follows a Poisson distribution.

$$N(t_3) - N(t_2) = \sum_{t=t_2+1}^{t_3-\Delta} x(t) + \sum_{t=t_3-\Delta+1}^{t_3} x(t) \sim Poisson((t_3 - \Delta - t_2)\mu + c\Delta\mu), \qquad (3.9)$$

A natural estimator of the final accrual size $N(t_3)$ can be calculated as the summation of the number of patients accrued up to time t_2, $N(t_2)$, and the estimated additional accrual from $t_2 + 1$ to the predicted closure date t_3:

$$\widehat{N(t_3)} = N(t_2) + \{(t_3 - \Delta - t_2) + c\Delta\}\widehat{\mu},$$

The standard deviation of the estimated total accrual can be calculated based on the Poisson distribution, $S.D.(\widehat{N(t_3)}) = [\{(t_3 - \Delta - t_2) + c\Delta\}\widehat{\mu}]^{\frac{1}{2}}$.

3.7 DETERMINATION OF t_1

The estimation of the accrual rate $\lambda(t)$ can be calculated based on two methods given the likelihood functions 3.2 and 3.3. As we discussed in previous section, if the accrual rate stays stable after a certain time point, it would be simpler to estimate the constant accrual rate $(t) = \mu$ based on likelihood function 3.3. The estimation of the accrual rate or the constant stabilized rate μ is critical because it affects the projection of the future closure date. The precise estimation of accrual rate is contingent on the appropriate choice of t_1, when the accrual rate becomes stable, and an important assumption that patient accruals during disjoint time intervals are independent. In this section we will focus on the appropriate choice of t_1 when the independence assumption stands.

There are different ways to determine t_1, when the accrual rate becomes stable. One way is based on model diagnosis after choosing t_1 based on the observed accrual pattern up to t_2, such as residual analysis, Kolmogorov-Smirnov (K-S) test, and likelihood ration statistic G^2.

For a pre-determined t_1, if the accrual rate is constant after t_1, the residuals of t_1 should follow a standard normal distribution. The residuals can be calculated as:

$$\epsilon(t) = \frac{x(t) - \widehat{\mu}_a}{\widehat{\mu}_a^{\frac{1}{2}}}, \quad t = t_1 + 1, \ldots, t_2. \tag{3.10}$$

Under a properly chose t_1, the residuals should behave as $N(0, 1)$. A plot of residual over time or a Q-Q plot should reveal whether such a choice is appropriate see how close it is from a the standard normal distribution.

A Kolmogorov-Smirnov test is also checking the distribution of residuals against the standard normal distribution. The test statistic can be carried out as:

$$S_{KS} \quad = \quad \max\{\epsilon(t) : t_1 + 1 \leq t \leq t_2\} \tag{3.11}$$

with the null hypothesis that the sample is drawn from the reference distribution for the one-side test. Here, the reference distribution is $N(0, 1)$.

Another approach is to use the likelihood-ratio statistic G^2 to test the goodness of fit. The test statistic is

$$G^2 = 2 \sum_{t=t_1+1}^{t_2} X(t) \log \frac{X(t)}{\widehat{\mu}}. \tag{3.12}$$

which follows a Chi-square distribution with a degree of freedom $(t_2 - t_1 - 1)$.

The largest p-values of the K-S test and G^2 statistical tests lead to the optimal choice of t_1.

Since the accruals from t_1 to t_2 follow a homogeneous Poisson process, we could also check the model assumption for the homogeneous Poisson process, which is the ratio of sample variance to mean should be close to 1 for an appropriate choice of t_1.

In practice, a possible range for t_1 can be chosen based on the raw accrual pattern. Then with t_1 varying in this pre-determined range, the above statistical tests were carried out to find out the optimal choice of t_1 that leads to the largest p-values for those test statistics and the ratio of sample variance to mean close to 1.

3.8 CORRELOGRAM FOR CHECKING SERIAL CORRELATION

After selecting an appropriate choice of t_1, we need to test another important assumption that patient accruals during disjoint time intervals are independent, which is also required for the precise estimation of accrual rate. For equally spaced time series data, the correlogram is a useful tool to detect whether there are serial correlations (Diggle, 1990 [8]). The daily accrual could be considered as equally space data. For the daily accrual data from a large multicenter trial, correlogram can be used to check whether there is serial correlation among adjacent time intervals. For the sequence of daily accrual data $\{x_t : t = 1, \ldots, n\}$, we use $\overline{x}$ to denote its sample mean, $\overline{x} = (\sum x_i)/n$, and define the kth sample autocovariance coefficient,

$$g_k = \sum_{t=k+1}^{n} (x_t - \overline{x})(y_{t-k} - \overline{y})/(n-k) \tag{3.13}$$

In the above formulae, the conventional use of the denominator is n instead. When $n >> k$, as is the case here, there is no essential difference. Then, the kth sample autocorrelation coefficient is

$$\gamma_k = g_k/g_0. \tag{3.14}$$

A plot of γ_k against k is called the correlogram of the data $\{x_i\}$. The plot of correlogram is to check whether there is evidence of any serial dependence in the data series. The values of γ_k greater than $2/\sqrt{n}$ in absolute value can be regarded as significant at about the 5% level.

4.0 APPLICATION ON TWO CLINICAL TRIALS FROM THE NSABP

4.1 APPLICATION BASED ON NSABP TRIAL B-38

To illustrate the proposed method, we consider the accrual data of the motivating study, the NSABP trial B-38, which was designed to compare the efficacy of three adjuvant chemotherapy regimen for node-positive breast cancer patients. More information on this study was introduced in Section 1.1. It was designed that a sample size 4,800 would achieve desired power for the primary hypothesis. More than a hundred of sites participated in the process of patient recruitment. The initiation date for each site was unknown. The first patient was enrolled on November 3, 2004. By the end of March 20, 2007, 4465 patients were randomized. The Biostatistical Center was required to make a projection on the trial closure date. Back then, a crude mechanism was considered in making the decision, which is the average daily accrual during the 3-month period before March 20, 2007 and added 3 more days. It was determined that the closure date was May 3, 2007. At the end of the designated closure date, a total of 4894 patients were accrued into B-38. Total of 94 patients were over accrued. Although several additional accruals might be necessary for this trial because dozens of patients withdrew consent or lost to follow-up later on. But accruing large number of additional patients often leads to unnecessary cost and financial burden on the trial sponsors.

Therefore, our purpose is to project a more precise closure date and avoid the additional accrued patients but make sure to accrue enough patients so that the study has sufficient power to make meaningful inferences. At first, we will use the proposed method to estimate the past accrual rate before March 20, 2007. Then, we will subsequently predict the closure date and the estimated total accrual by the end of the closure date.

4.1.1 Predication of closure date with t_1 fixed

The main reason that we need hundreds of institutions or sites to participate in the patient accrual is because there is a small chance for a single site to put a patient on the trial during a short period, say a work day. We assume that the accrual from different sites is independent among each other and the accrual during a certain time period is independent from another disjoint time period. Therefore, the daily accrual rate follows a Poisson process. Based on the observed accrual in Figure 1, at the beginning of the recruitment, the accrual rate was relatively slow and non-linear. After several months, the accrual rates became stable. However, during the final few days, the accrual rate is much higher than the past accrual rates. Therefore, it is reasonable to use the Poisson process model defined by 3.1, to estimate the accrual rate $\lambda(t) = \mu$, where t indicates the accrual days. After excluding weekends and holidays, if the first patient was accrued on November 3, 2004 (Wednesday) which is denoted as day 1, and November 8, 2004 (Monday) would be day 4. Then, March 20, 2007 when the coordinate data center made the prediction of closure date would be day 594 ($t_2 = 594$).

With observed accrual up to March 20, 2007, we need to predict the future closure date T when the total accrual would exceed the target sample size $N = 4800$. First, we need to determine t_1 when the accrual rate became stable. Based on the observed accrual in Figure 1, the accrual rate during about the first 4 months, from November 2004 to February 2005, was not stable and non-linear. As a starting point, we could arbitrarily choose a date for t_1 and assume that the accrual rate became stabilized on that day. Based on model diagnosis tools, we could evaluate if the choice of t_1 is appropriate. The methods of model diagnosis, residuals analysis, Kolmogorov-Smirnov test, and likelihood ration statistic, were discussed in Section 3.7.

For example, we could first choose t_1 on March 7, 2005 ($t_1 = 83$). Follow the method in Section 3.3, the MLE of the constant daily accrual rate after March 7, 2005 is $\widehat{\mu}$=8.13 and its standard error is $se(\widehat{\mu})$=0.126. However, the choice of t_1 would affect the estimate of daily accrual rate, which will be discussed in next section.

From the overall observed accrual data in B-38, as shown in Figure 1, the daily accrual rate during the last 5 days was much higher than the daily accrual rate during the middle of

the accrual period. This is because the participating institutions tried to enroll their patients into this trial since they believe their patients would be benefited. In 3.1, the accrual rate during that 5-day period is presented as $c\mu$, where c is a constant. In practice, we would have to estimate this information using accrual data from similar historical trials, in term of disease status, treatment regimen, coordinate center, and participant institutions. However, it is always possible that such a similar completed clinical trial is not available. In that case, we could conduct some sensitivity analysis to help determining c. It will be discussed in later section. Here, we calculated the MLE for c based on the overall accrual data from B-38, in order to illustrate the method discussed in Section 3.4 on how to get MLE of c. $\hat{c} = 2.24$ is estimated based on the overall accrual data from B-38 given t_1=83 (7MAR2005).

Using the proposed method in Section 3.5, there are two methods to determine the future closure date T. One is the average of time $E(T)$ when the overall accrual would pass the target sample size N. Another one is the quantile of T, which is the probability of the total accrual can pass the sample size N when $t_\alpha = min\{t : pr[T \leq t] \geq 1 - \alpha\}$. In the following context, α was chosen as 0.05. In order to get the expectation of T, we need to first calculate the estimated average of additional accrual days and its standard error, which are $\widehat{E(W)}$=26 and $se(\widehat{E(W)})$=0.487, respectively. The projected total closure days $\widehat{E(T)}$ is calculated as $\widehat{E(W)} + t_2$, where $t_2 = 594$. The projected closure date $\widehat{E(T)}$ and the predicted total accrual by the end of the estimated closure date $E[N(t)]|_{t=\widehat{E(T)}}$ are presented in Table 1.

The results of $\widehat{t_\alpha}$, its 95% CI and the estimated total number of final accrued patients $E[N(t)]|_{t=\hat{t_\alpha}}$ are also presented in Table 1. Based on the results in Table 1, it is more conservative for predicting the closure date to use the quantiles of T. The results by this method requires 3 additional accrual days, which allows more patients to be accrued. Its 95% confidence interval (CI) is estimated by the bootstrap method.

The accrual days, $\widehat{E(W)}$, $\widehat{E(T)}$ and $\widehat{t_\alpha}$ were always rounded to their ceilings, that is, a value of 21.2 would be rounded to 22. However, when we calculate the expected number of total accrual by the closure date, $E[N(t)]|_{t=\widehat{E(T)}}$ or $E[N(t)]|_{t=\widehat{t}_\alpha}$, the floor of the estimates was used, i.e., 4821.8 would be reported as 4821.

In Appendix, Table 10 shows the actual date and accruals towards the end of the recruitment. We can always compare our results with the real data. At the end of the day

Table 1: Predicted closure date and estimated total final accrual at t_1=7MAR2005 in B-38

$\widehat{E(T)}(SE)$	$E[N(t)]\|_{t=\widehat{E(T)}}$ (SD)	$\hat{T}_{0.05}$ (95% CI)	$E[N(t)]\|_{t=\hat{T}_{0.05}}$ (SD)
620 (0.487)	4805 (18.5)	623 (622, 624)	4831 (19.1)

620, total of 4842 patients were actually enrolled. And, at the end of the day 623, total of 4872 patients were accrued. The actual total accruals were calculated as the observed total accrual during the last five days prior to May 3, 2007, and the period between $t_2 + 1$ and $t_3 - 5$. This is because that we have to always consider the possibility of high accrual at the end of the recruitment. It has been observed in many large multi-center phase III trials. It means that if the trial close on day 620, a high accrual period between days 616 to 620 will be expected. The results of the actual total accrual do not match the results from the proposed method very well. It can be one of the indicator that the choice of t_1 is not precise. The results will be compared with other different choices of t_1.

Given that t_1=83 (7MAR2005) and $\hat{c} = 2.24$, the results in Table 1 can be interpreted as: in average the overall accrual would surpass 4800 during day 620. In average, a total of 4805 patients would be accrued by the end of the day with a standard deviation of 18.5. We have 95% confidence that the overall accrual would surpass 4800 by the end of day 623. If day 623 is the designated closure date, in average a total of 4831 patients would be accrued with a standard deviation of 19.1. We could also change the value of α in order to get different quantiles of T.

The residuals resulted from fitting a homogeneous Poisson process on accrual data during t_1 and t_2 are plotted in Figure 2. The Q-Q plot of the residuals of the estimated accrual rate $\widehat{\mu}$ is presented in Figure 3. It seems that the distribution residuals skewed a little to the right, compared with a standard normal distribution. The results of Kolmogorow-Smirnov test and likelihood-ratio statistics are $S_{KS} = 0.0992$ (p value< 0.001) and $G^2 = 635.11$ with df=510 (p value< 0.001), respectively. This suggests that accrual during March 7, 2005 and March 20, 2007 did not follow a homogeneous Poisson process. We could observe that the

most of the residuals during the first few days are below 0, and the values of S_{KS} and G^2 need to be compared with the results based on different choices of t_1. In next section, we will explore the effects of t_1 on the estimates of the accrual rate and the projection of the closure date.

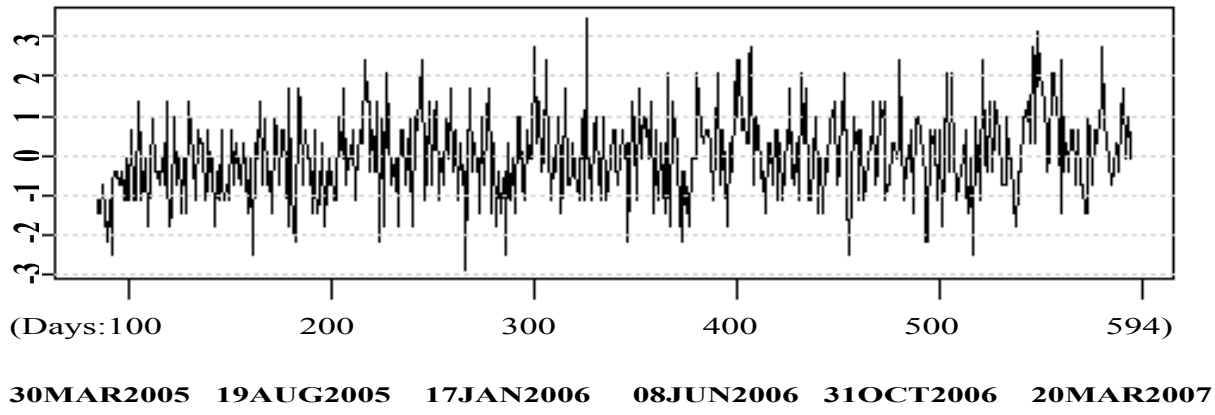

Figure 2: Residuals of accrual rate over time at t_1=07MAR2005 in B-38.

4.1.2 The choice of t_1

The precise estimation of accrual rate is contingent on the appropriate choice of t_1. The method of choosing an appropriate t_1 and the parameter estimates based on selected t_1 will be discussed here.

Based on the observed accrual from t_0 to t_2 in Figure 1, t_1 was chosen on March 7, 2005 in above section. However, the Q-Q plot shows a majority of the residuals during the early days are negative. It means the daily accrual rates are relatively lower at the first several days after March 7, 2005 compare to later days. This suggests that t_1 should be chosen at a latter date. In this section, we propose a simple algorithm to help choose t_1 within a predetermined time period.

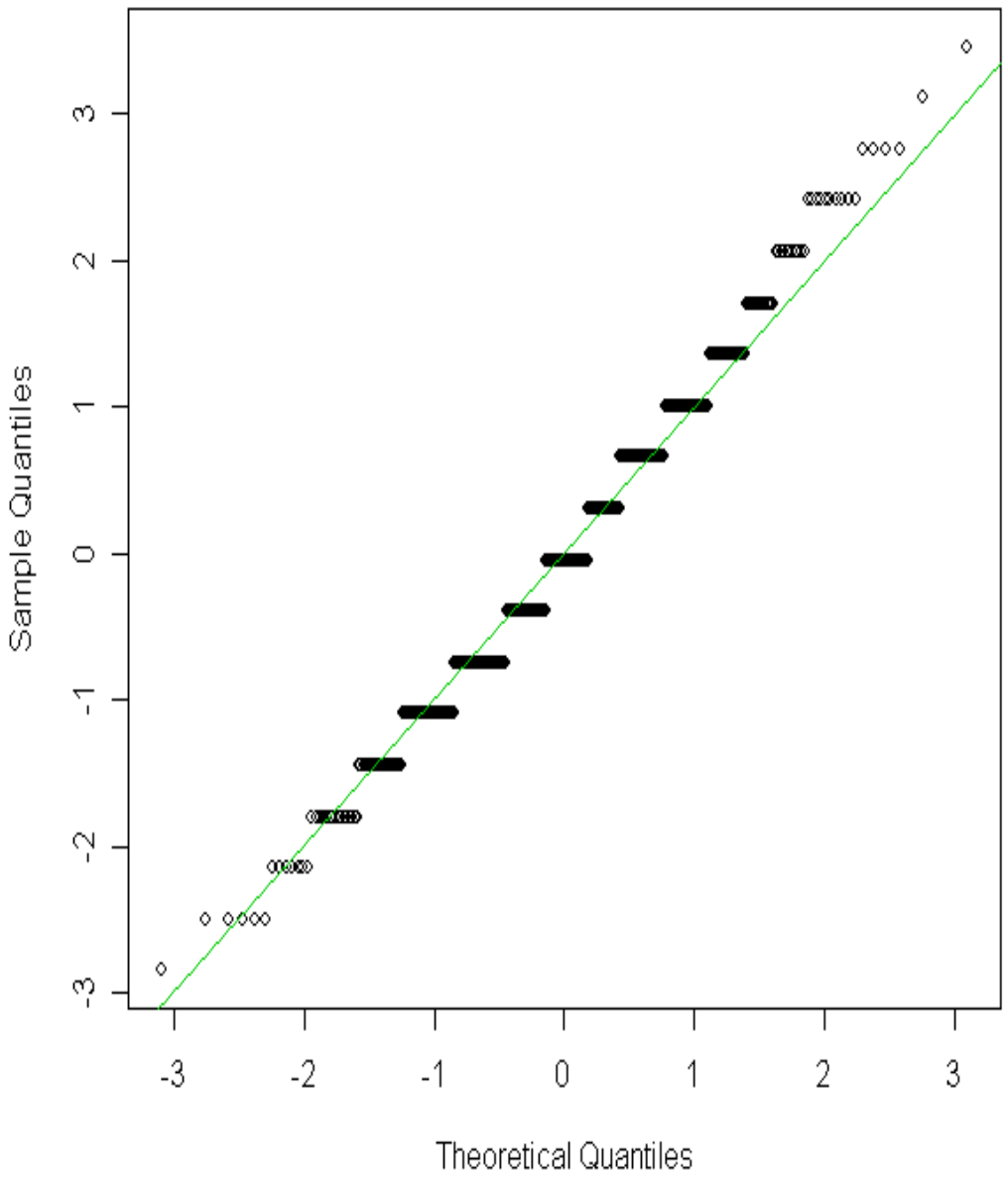

Figure 3: Q-Q plot of the residuals of accrual rate at t_1=07MAR2005 in B-38.

First, based on the overall observed accrual pattern from t_0 to t_2, we could determine a potential time period from which t_1 could be chosen. Second, we calculate the ratios of sample variance to mean and p-values of G^2 for each choice of t_1 within the predetermined time period. From t_1 to t_2, the accruals should follow a homogeneous Poisson process. Therefore,

the ratio of sample variance to mean should be relatively close to 1 for an appropriate choice of t_1. When the p-value of G^2 is also relatively large, it indicates a precise choice of t_1. Then, we generate a figure, which includes both the information of the ratios and p-values given t_1 varying in the predetermined time period. It helps to visualize the results and select t_1.

Based on the accrual pattern from B-38 and the model fitting information from above section given t_1=7Mar2005, we decided that the potential time period for t_1 to choose from is March 30, 2005 to February 20, 2007 which was 20 days ahead of t_2.

After calculating the ratios of sample variance to mean and p-values of G^2 for various choices of t_1, we generated a Figure 4 including both the information of the ratios and p-values given t_1 varying in the predetermined time period. It shows in Figure 4 that the ratio of sample variance to mean is relatively close to 1 and the p-value of G^2 is relatively large towards the end of 2006, ratio=1.01, p-value=0.429. So, the optimal choice of t_1 was December 29, 2006 for this trial.

Given t_1=29DEC2006 and $c = 2.24$, $\widehat{\mu} = 9.80$, $se(\widehat{\mu}) = 0.426$, the p-value for K-S test is 0.536 and $S_{KS} = 0.1059$. The predicted closure date and estimated total final accrual are presented in Table 2. The plot of residuals over time and the Q-Q plot of residuals are presented in Figure 5 and 6. Compare to the results of K-S test and the plots based on t_1=7MAR2005, the fit to the Poisson model was dramatically improved. However, the standard error of $\widehat{\mu}$ increased because that the number of observed data points decreased as t_1 was chosen closer to t_2. Also, the absolute values of residuals are greater or equal than 3 in absolute value should be considered as outliers. There are not outliers observed from Figure 5. If observed any, the outliers should be taken out and the accrual rate needs to be re-estimated to compare with the original results, in order to determine the influence of the outliers.

In Appendix, Table 10 shows the actual date and accruals towards the end of the recruitment. Compare our results in Table 4 with the real data in Table 10, by the end of day 616, in average 4809 patients were expected to enroll and 4802 patients were actually enrolled. At the end of the day 618, in average 4829 patients were expected to enroll and 4819 patients were actually accrued. The results of the actual total accrual match the results from the proposed method pretty well. It indicates that the choice of t_1=29DECf2005 is appropriate.

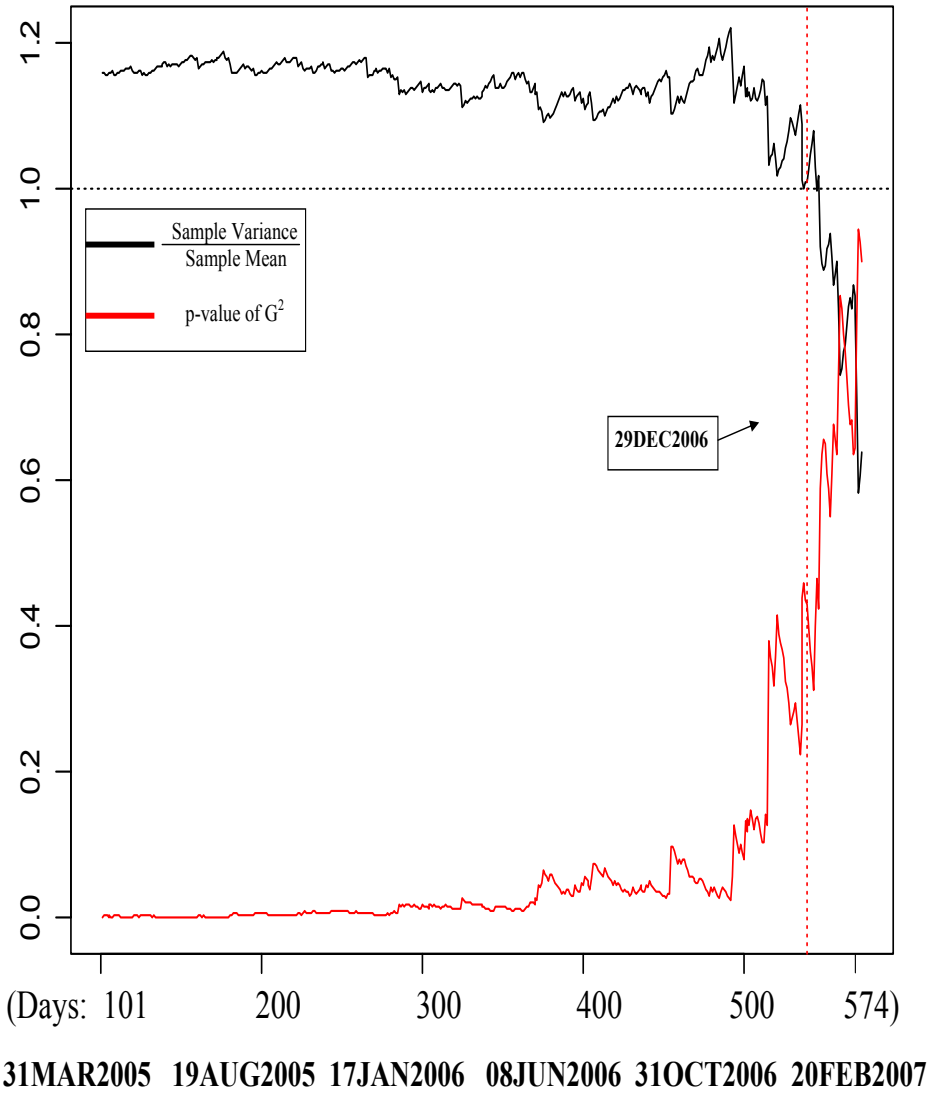

Figure 4: The ratios of sample variance to mean and p-values of G^2 for various choices of t_1 in B-38.

4.1.3 When the constant c is unknown

After a closure date is predicted, most of the clinical trials usually experience a dramatic increment in the daily accrual rate during the last few days before the closure date. It is

Table 2: Predicted closure date and estimated total final accrual at t_1=29DEC2006 in B-38

t_1	$\widehat{E(T)}$ (SE)	$E[N(t)]\|_{t=\widehat{E(T)}}$ (SD)	$\hat{T}_{0.05}$ (95% CI)	$E[N(t)]\|_{t=\hat{T}_{0.05}}$ (SD)
29DEC2006	616 (2.657)	4809 (16.3)	618 (616, 621)	4829 (16.9)

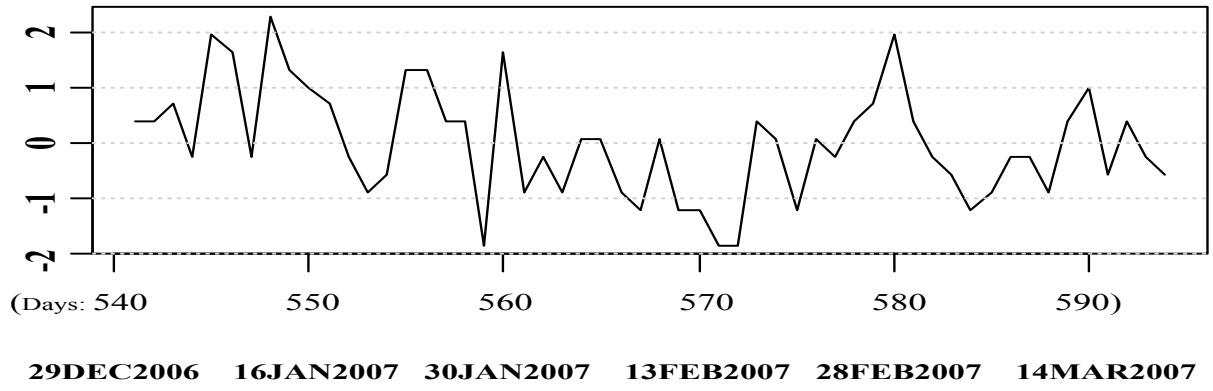

Figure 5: Residuals of accrual rate over time at t_1=29DEC2006 in B-38.

because the participating sites try to put their patients on the trial before the deadline since they believe that their patients would be benefited from the clinical trial. The accrual rate for the last few days is estimated as $\lambda(t) = c\mu$, as in 3.1. The constant $c > 1$ could be speculated from the accrual patterns of past multi-center trials carried out by the same coordinating center, which share some similarities, such as patient characteristics and treatment regimen. This information is important to know since it affects the prediction of closure date and the estimated total accrual.

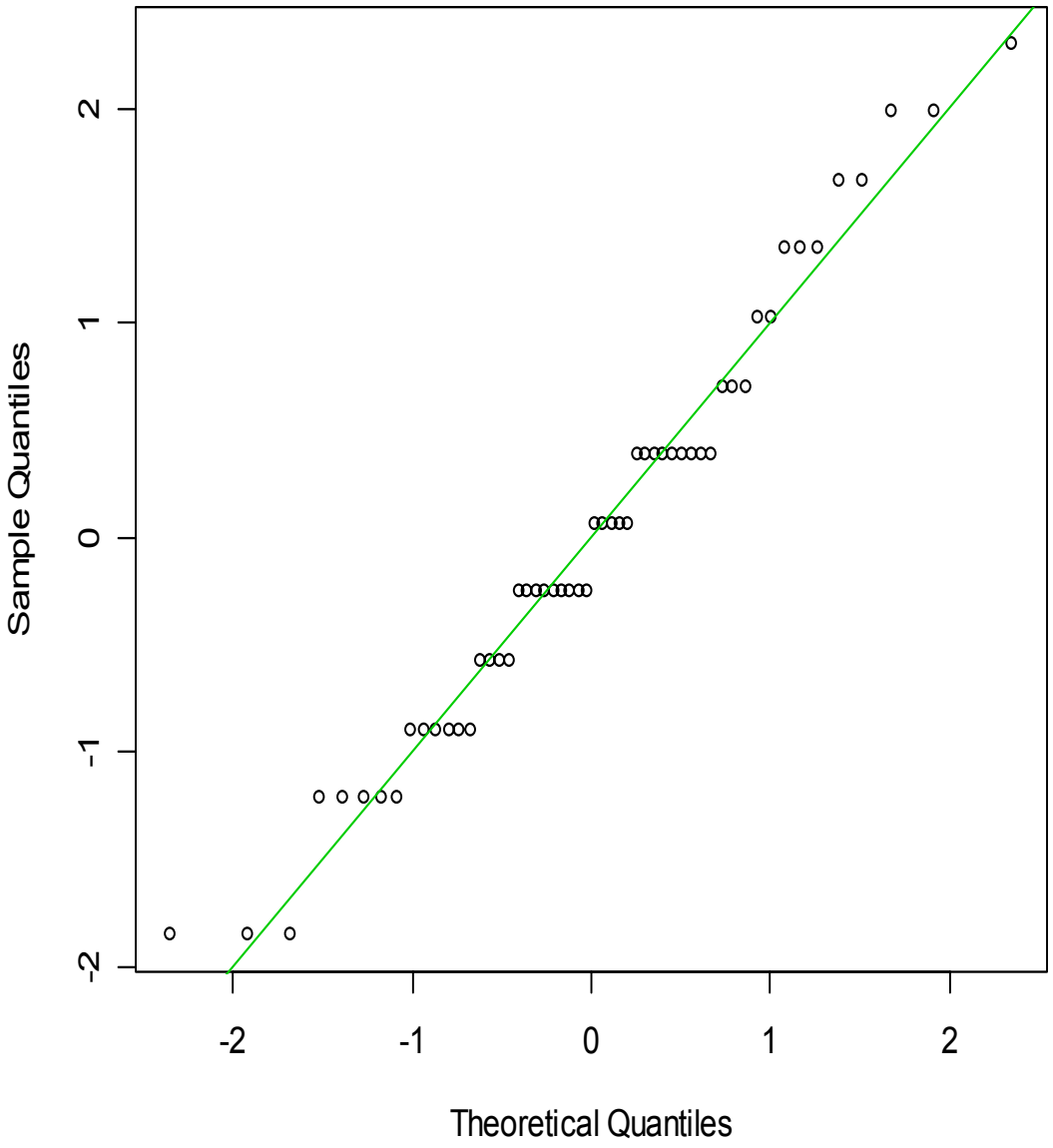

Figure 6: Q-Q plot of the residuals of accrual rate at t_1=29DEC2006 in B-38.

However, it is possible the information of c is unavailable for some clinical trials if there are no similar completed historical clinical trials available. In this section, we will conduct sensitivity analysis to evaluate the influence of c when we set c at different values, and also compare the results when c is fixed and the choice of t_1 is varied. When c=1, we assume

there is no increased accrual rate during the last few days before the closure date. When c is increasing, the accrual rate increases at the last few days. Based on the overall B-38 accrual data, we know c is around 2.2. So, we will choose c at 1.5, 2, and 2.5 and compare the results of predicted closure date and estimate total accrual.

In Table 4, t_1 was chosen on December 29, 2006 based on the conclusion from the model diagnosis in last section. In Table 3, t_1 was chosen about 2 weeks before December 29, 2006, which is December 14, 2006. The results from Tables 3 and 4 show that when c is smaller it takes extra days to complete the accrual given any choice of t_1. For example, when c=1.5 it takes almost additional 5 days as comparing to c=2.5 for t_1 on December 14 or 29 in 2006. However, when we settle with a value of c, the predicted accrual days does not vary much, about 1-2 days difference as t_1 varies in about 2 weeks from December 14, 2006 to December 29, 2006. It requires additional 4-5 days if t_1 was chosen on March 7, 2005, more than 1.5 years ago from December 29, 2006, see Table 1 and Table 2, when the value of c is the same.

After getting all of the results, we need to present the results to investigators or researchers. It is important to show all of the results based on varied t_1 and c, and then help investigators understand the results. For example, the results in Table 2 can be interpreted as: given t_1=540 (29DEC2006) and $c = 2.2$, the results of $\widehat{E(T)}, (SE)$ and $E[N(t)]|_{t=\widehat{E(T)}}, (SD)$ can be interpreted as: in average the overall accrual would take total 616 days with a standard error of 2.66 and in average a total of 4809 patients could be accrued by the end of that day with a standard deviation of 16.3. For the results of $\hat{T}_{0.05}, (95\%CI)$ and $E[N(t)]|_{t=\hat{T}_{0.05}}, (SD)$, it can be interpreted as: we have 95% confidence that the overall accrual would pass 4800 by the end of day 618. If day 618 is the designated closure day, in average there will be total 4829 patients accrued with a standard deviation of 19.1. In Appendix, Table 10, we can find out that day 616 is April 19, 2007 and day 618 is April 23, 2007.

Also, it's important to help them understand and select an appropriate c given the sensitivity analysis. We could start with asking the following questions:

(1) After the closure date disseminated, would you expect the same accrual pattern as what we observed in the past?

(2) If the accrual rate would increase during the final few days, how long is the window and how much increment would be expected?

Table 3: Predicted closure date and estimated total final accrual at t_1=14DEC2006 for varying c in B-38

c	$\widehat{E(T)}$ (SE)	$E[N(t)]\|_{t=\widehat{E(T)}}$ (SD)	$\hat{T}_{0.05}$ (95% CI)	$E[N(t)]\|_{t=\hat{T}_{0.05}}$ (SD)
1.5	620 (2.93)	4810 (16.3)	622 (620, 625)	4829 (16.9)
2	617 (2.93)	4806 (16.2)	620 (617, 622)	4834 (17.0)
2.5	615 (2.93)	4810 (16.3)	617 (615, 620)	4829 (16.9)

Based on the answer of the first question, we could decide if there are any potential increased or maybe even decreased daily accrual rates at the end of the recruitment for some cases. Then, we could decide what the value of c could be. If the investigator expects an increment at the last a few days of the recruitment, we will consider $c > 1$. If the increment is not expected, then c is most likely close to 1. If the investigator expects a very slow accrual towards the end and thinks the patient enrollment can get sloppy at the end, we need to remind them about the phenomenon we observed based on NSABP large multicenter phase III cancer trials. If they think their trials are different than what we usually conducted at NSABP and certain about a slow accrual at the end of the accrual, $0 < c < 1$ need to be considered. The second question would help us to decide the value of Δ, which is the number of days at the end of the accrual with increased rates. Besides, it also asks that how much of the increment would be expected. It will help us to decide if c is more towards 1 for a small increment or 2 or 3 which associated with a larger increased rates. However, based on what we observed from the trials from NSABP, we do not suggest a very large value of c, no larger than 2.5 or 3. When c is less or equal than 2.5, it provides a more conservative projection of the closure date. As we discussed previously, a large value of c could lead to a shorter accrual period and less accrued patients than the target sample size. At the end, after presenting the predicted accrual dates ($\widehat{E(T)}$ or $\widehat{t}_\alpha$) and the total predicted accruals to the investigators under the various plausible choices of c, the investigators could make their decision based on their confidence, the available resource and other considerations.

Table 4: Predicted closure date and estimated total final accrual at t_1=29DEC2006 for varying c in B-38

c	$\widehat{E(T)}(SE)$	$E[N(t)]\|_{t=\widehat{E(T)}}$ (SD)	$\hat{T}_{0.05}$ (95% CI)	$E[N(t)]\|_{t=\hat{T}_{0.05}}$ (SD)
1.5	619 (2.66)	4814 (16.4)	621 (618, 623)	4833 (17.0)
2	616 (2.66)	4809 (16.3)	618 (616, 621)	4829 (16.9)
2.5	614 (2.66)	4814 (16.4)	616 (614, 618)	4933 (17.0)

4.1.4 Checking for Serial Correlation

Accrual data from multicenter trials can be treated as an equally spaced series. Correlogram is helpful to check whether there is serial correlation among accruals during adjacent intervals (Diggle, 1990). The kth sample autocovariance coefficient g_k is calculated as in 3.13. The kth sample autocorrelation coefficient γ_k is calculated as in 3.14. The overall correlogram is a plot of γ_k against k. Here, we choose k=30, for 30 days starting from t1. The plot of correlogram is to check if there is evidence of any serial dependence in the data series. The values of γ_k greater than $2/\sqrt{n}$ in absolute value can be regarded as significant at about the 5% level, where $n = t_2 - t_1$. At t_1=29DEC2006 (days=540) and t_2= 20MAR2007(days=594), Figure 7 suggests none or very small serial correlation among this accrual data since there is only one observation's autocorrelation coefficient γ_k is a little bit over 0.27 on January 1, 2007 and majority of the γ_k's are within the range of ± 0.27. The 0.27 is calculated as $2/\sqrt{n}$, where $n = 594 - 541 = 54$.

4.2 APPLICATION BASED ON TRIAL NSABP C-08

In NSABP trial C-08, a phase III, randomized clinical trial to compare disease-free survival of patients with resected stage II or III adenocarcinoma of the colon treated with two adjuvant

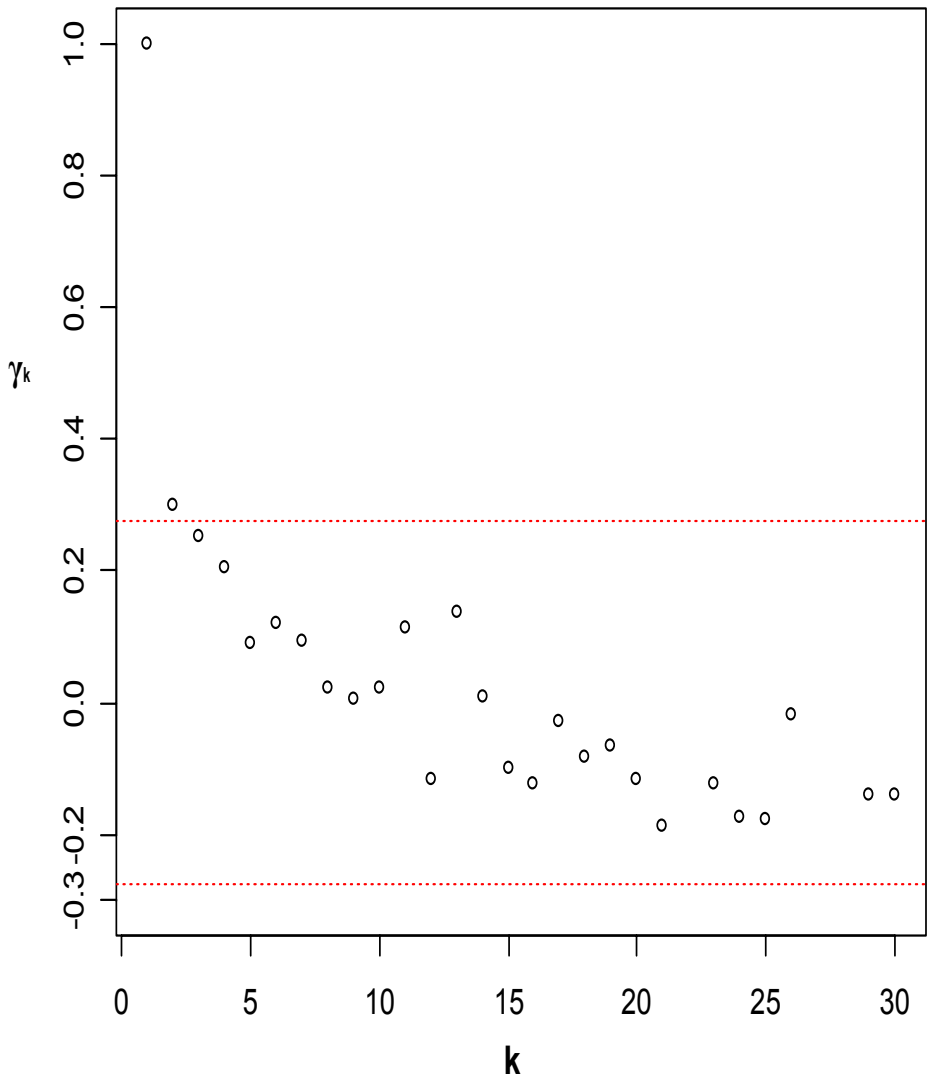

Figure 7: Overall correlogram for B-38, t_1=29DEC2006.

chemotherapy regimen: luorouracil, leucovorin calcium, and oxaliplatin with versus without bevacizumab. The pre-determined sample size was 2632. The trial was opened on September 15, 2004 and the first patient was accrued on October 14, 2004. On August 23, 2006 the coordinating data center was required to project an accrual closure date for this trial based on the observed accrual pattern and send it to participating sites several weeks in advance. Based on the average daily accrual during the 3-month period before August 23, 2006 and added 3 more days, October 6, 2006 was determined as the closure date. At the end of the determined closure date, 2710 patients were accrued and randomized in this trial and a total of additional 78 patients were accrued.

Figure 8 shows the daily accrual in trial C-08 over time, which illustrates the same phenomenon as it observed in Figure 1 for trial B-38. During the first 3-4 months, the accrual rate was increasing steadily and the accrual pattern became relatively stable after March 2005. Again, it reflects the fact that it took time for the participating sites to meet regulatory requirements and start recruiting patients for the trial. So, we choose t_1 on March 10, 2005 (t_1=100) as a starting point and assume the accrual rate became stabilized on that day. t_2 is August 23, 2006 (t_2=467), which is the same as the actual data, when the accrual pattern was observed by and the future closure date was determined on. The MLE of the daily accrual rate is $\widehat{\mu} = 5.89$ and its standard error is $se(\widehat{\mu}) = 0.127$.

From the overall observed accrual data in trial C-08, as shown in Figure 8, the daily accrual rate during the last 5 days was much higher than the daily accrual rate before that period. It increased about twice of the average daily accrual during the middle of the accrual period. This is because the participating sites tried to enroll their patients into the trial since they believed that the trial would benefit their patients. As it introduced in (3.1), the accrual rate during the last 5-day is estimated as $c\mu$. c could be speculated based on the accrual pattern from a similar completed historical clinical trial, which was carried out by the same coordinating center and has the same patient characteristics and treatment regimen as it utilized in trial C-08. Here, $\hat{c} = 1.8$ was estimated using the overall accrual from October 14, 2004 to October 6, 2006 based on the accruals in C-08. It is just to illustrate the calculation of c. When historical trials were not available, the sensitivity analysis of c will need to be conducted and presented later.

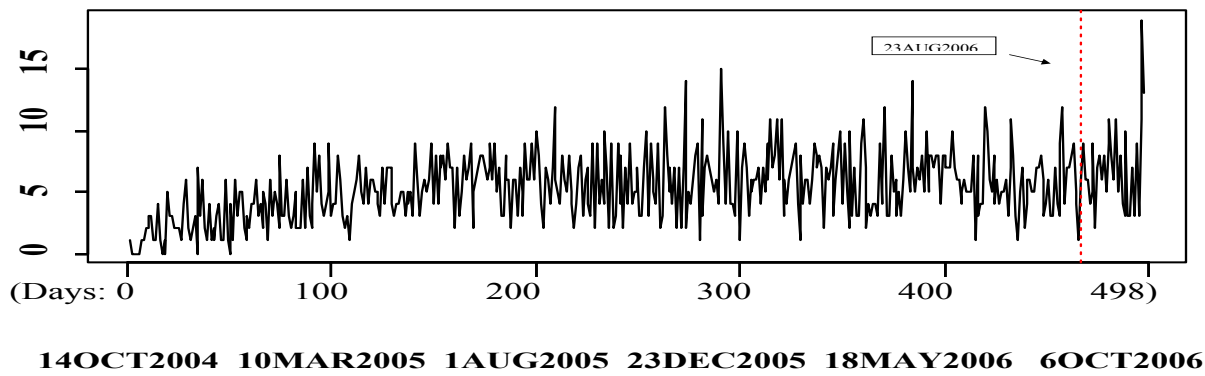

Figure 8: Overall daily accrual from NSABP trial C-08.

Given $t_1 = 100$ (10MAR2005) and $c = 1.8$, the estimated average of additional accrual days and its standard error are $\widehat{E(W)} = 22$ and $se(\widehat{E(W)}) = 4.23$. The projected total closure day is $\widehat{E(T)} = 489$, which is $t_2 + E(W)$. The predicted total accrual by the estimated closure date is $E[N(t)]|_{t=\widehat{E(T)}} = 2639$ with a standard deviation of 12.5, which could be interpreted as at the end of the day 489 (25SEP2006) the overall accrual will surpass the target sample size 2632 with a standard deviation of 12.5.

With $\alpha = 0.05$, $\hat{t}_\alpha = 492$ and it's 95% CI is (491, 493). The projected total closure days is $E[N(t)]|_{t=\hat{t}_\alpha} = 2657$ and its standard deviation is 13.1. Based on the results, we could predict that we have 95% confidence that the overall accrual would surpass 2632 by the end of the day 492 (28SEP2006). If day 492 is the determined closure date, in average a total of 2657 patients would be accrued with a standard deviation 13.1. The results are also summarized in Table 5.

The results of K-S test and Likelihood ratio statistics are $S_{KS} = 0.1066$ (p-value< 0.001) and $G^2 = 393.89$ with df=366 (p-value=0.151), respectively. The residuals resulted from

Table 5: Predicted closure date and estimated total final accrual at t_1=10MAR2005 in C-08

t_1	$\widehat{E(T)}$ (SE)	$E[N(t)]\|_{t=\widehat{E(T)}}$ (SD)	$\hat{T}_{0.05}$ (95% CI)	$E[N(t)]\|_{t=\hat{T}_{0.05}}$ (SD)
10MAR2005	488 (4.23)	2639 (12.5)	492 (491, 493)	2657 (13.1)

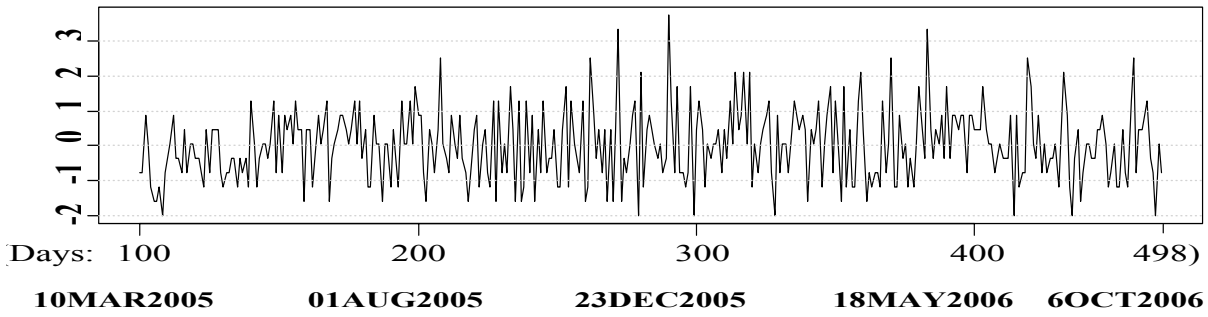

Figure 9: Residuals of accrual rate over time at t_1=10MAR2005 in C-08.

fitting a homogeneous Poisson process on accrual data during t_1 to t_2 are plotted in Figure 9. The Q-Q plot of the residuals of the estimated accrual rate $\hat{\mu}$ is presented in Figure 10. It seems that the distribution residuals skewed a little to the left, as comparing with a standard normal distribution. We could observe some of the residuals during the first and last few days are below 0 and potential outliers between October 2005 to May 2006 with relatively large residuals which are greater than 3. It indicates that a better t_1 needs to be determined based on the observed accrual pattern.

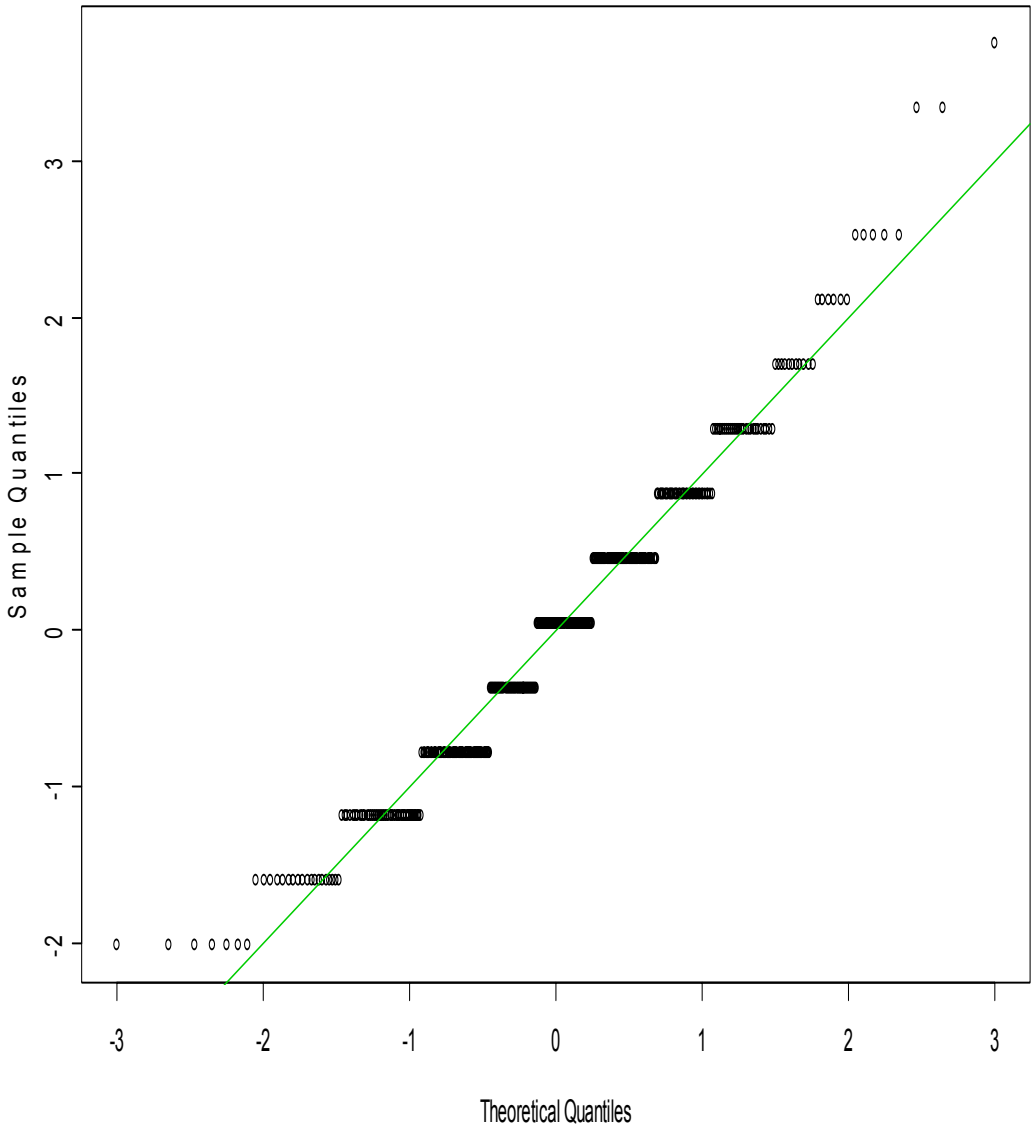

Figure 10: Q-Q plot of the residuals of accrual rate at t_1=10MAR2005 in C-08.

We follow the same algorithm introduced in Section 4.1.2 to select an appropriate time point for t_1. First, need to decide on the time interval for selecting t_1. Since the model diagnoses given t_1 on March 10, 2005 indicate t_1 needs to be chosen at a later time, we choose March 31, 2005 to July 26, 2006 as the predetermined time period from which to select t_1.

Then, we calculate the ratios of sample variance to mean and p-values of G^2 given t_1 varying in the predetermined time period. Figure 11 displays both of the information, the ratios of sample variance to mean and p-values of G^2, given t_1 varying within the predetermined time period. On April 26 2006, the ratio of sample variance to mean is close to 1 and the p-value of G^2 is relatively large (ratio=0.898, p-value=0.591), which indicates an appropriate choice of t_1. Therefore, t_1 is chosen on April 26, 2006 (t_1=384).

We still choose c=1.8, so we can compare the results with t_1 choosing at different date. The estimated accrual rate $\hat{\mu}$ is 5.99 and its standard error is 0.269. The projected total closure days and the predicted final total accruals are presented in Table 6. Comparing to $t_1 = 384$ (26APR2006), the estimated accrual rate stays the same around 6 patients per day but its standard error almost doubled and increases from 0.127 to 0.269. The increment standard error is because when $t_1 = 384$, the observed data points from t_1 to t_2 decreased. Hence the information contributes to the likelihood function for estimating μ decreased. The estimated total closure days and the predicted final total accruals do not varied much when c is the same.

The goodness of fit criteria for estimating the accrual rate given $t_1 = 384$ (26APR2006) improved compare to $t_1 = 100$ (10MAR2005). The results of K-S test and likehlihood-ratio statistics are $S_{KS} = 0.2767$ (p-value=0.277) and $G^2 = 78.44$ (p-value=0.591) with df=82. The residual plot and Q-Q plot (Figure 12 and 13) are also improved.

After determining the date of t_1, we need to conduct sensitivity analysis for varied values of c. Also, the results will need to be evaluated for the same value of c but given different date of t_1. t_1 was chosen on April 26, 2006 and April 13, 2006, about 10 days early from April 26, 2006, with c varying, Table 7 and 8 presented the results of the projected total closure days and the predicted total accrual by the end of the closure date. The same conclusion can be reached as in trial B-38. When c is smaller it takes extra days to complete the accrual, e.g. when c=1 it takes almost additional 5 days as comparing to c=2. However, when we settle with a value of c, the predicted accrual days does not vary much, about 1-2 days difference as t1 varies about 10 days apart.

The plot of correlogram is to check if there is evidence of any serial dependence in the data series. The correlogram for trial C-08, given $t_1 = 384$ (26APR2006) is presented in

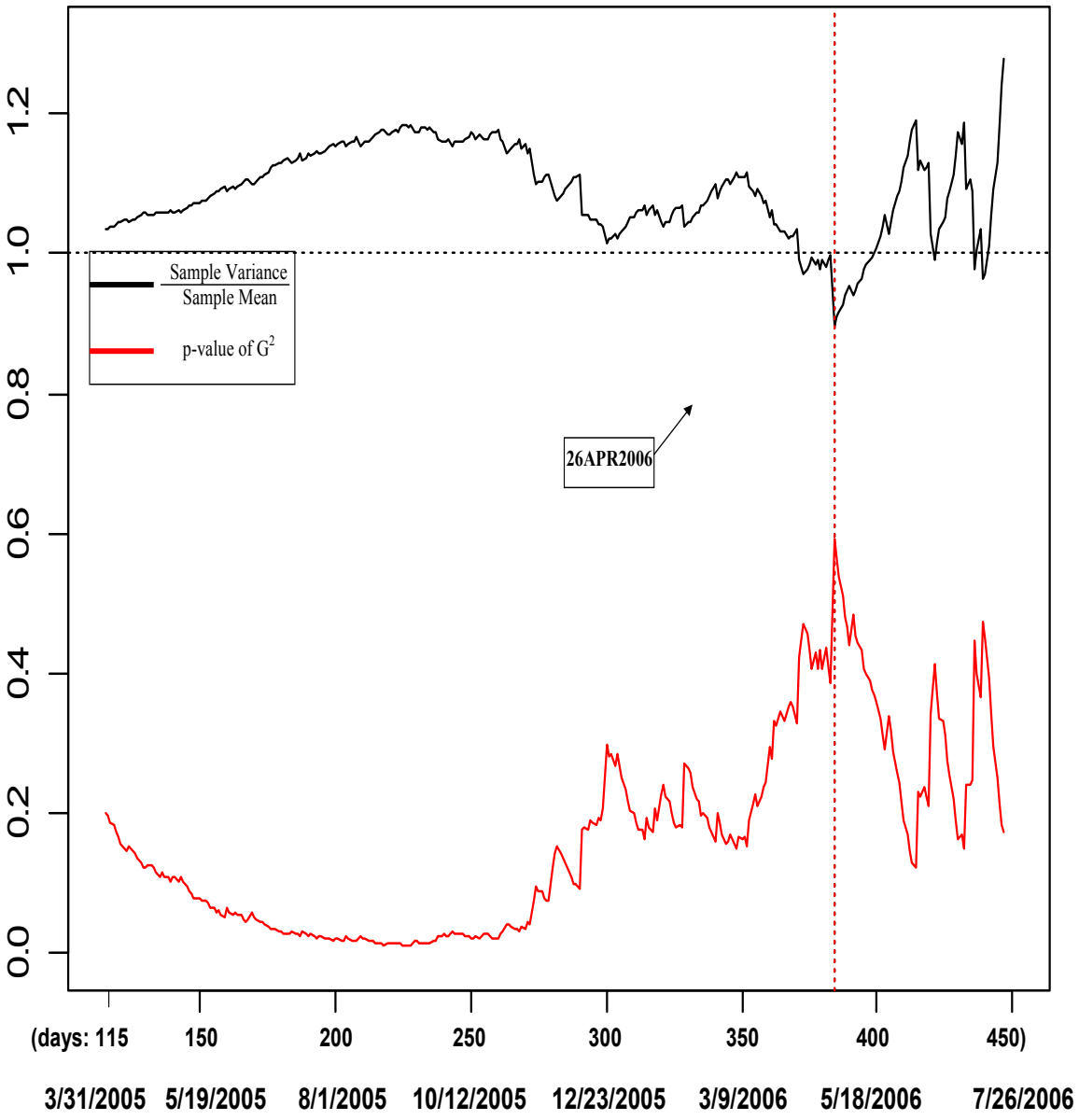

Figure 11: The ratios of sample variance to mean and p-values of G^2 for various choices of t_1 in C-08.

Figure 14. In the figure, we choose k=30, for 30 days starting from t1. The values of γ_k greater than $2/\sqrt{n} = 0.22$ in absolute value can be regarded as significant at about the 5% level, where $n = t_2 - t_1 = 467 - 384 = 83$. There are two data points greater than 0.22. It

Table 6: Predicted closure date and estimated total final accrual at t_1=26APR2006 in C-08

t_1	$\widehat{E(T)}$ (SE)	$E[N(t)]\|_{t=\widehat{E(T)}}$ (SD)	$\hat{T}_{0.05}$ (95% CI)	$E[N(t)]\|_{t=\hat{T}_{0.05}}$ (SD)
26APR2006	488 (4.10)	2640 (12.5)	491 (489, 494)	2652 (13.0)

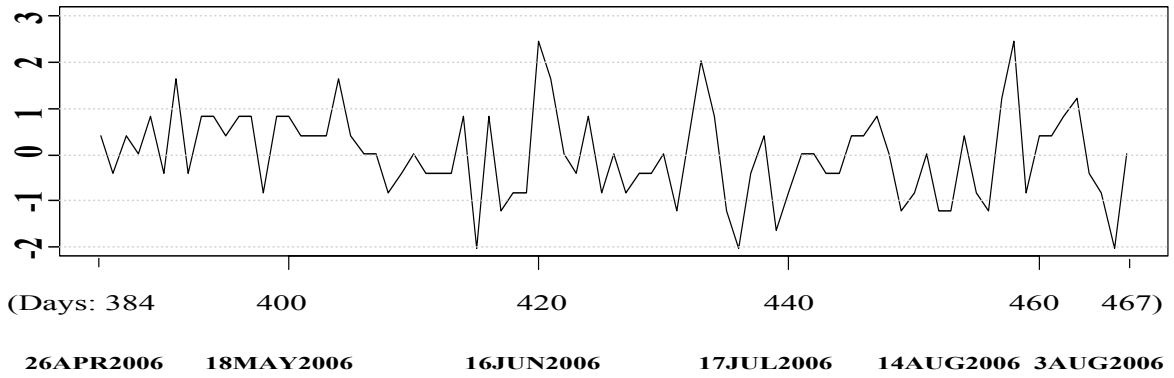

Figure 12: Residuals of accrual rate over time at t_1=26APR2006 in C-08.

shows very small serial correlation among this accrual data.

In conclusion, the results of C-08 are consistent with what we found in B-38. After selecting a time point t_1, when the accrual rates become stable, we need to either find c in a similar historical trial or conduct sensitivity analysis for varied c. Then, present the predicted future closure date and estimated total accruals to the investigators. They could make the decision based on their knowledge of the trial.

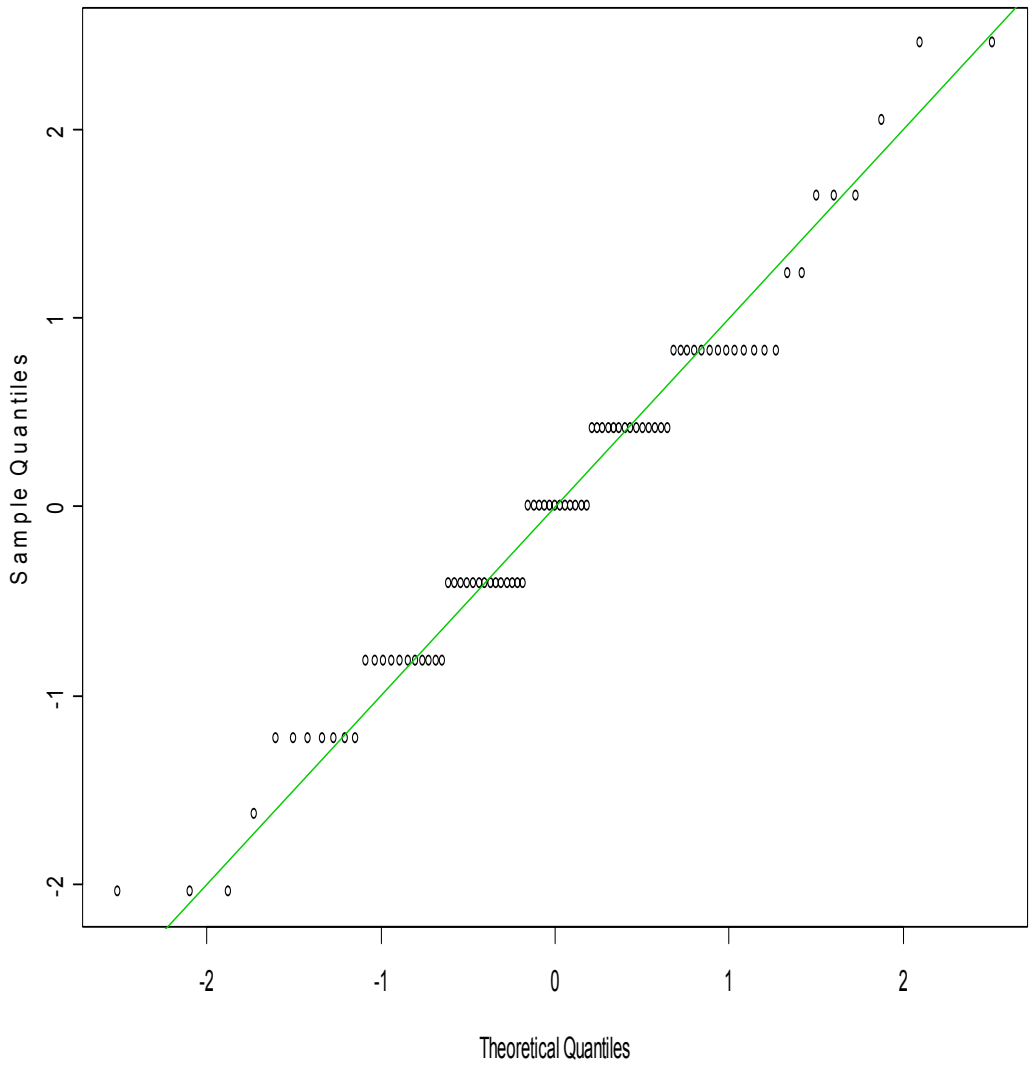

Figure 13: Q-Q plot of the residuals of accrual rate at t_1=26APR2006 in C-08.

Table 7: Predicted closure date and estimated total final accrual at t_1=13APR2006 for varying c in C-08

c	$\widehat{E(T)}$ (SE)	$E[N(t)]\|_{t=\widehat{E(T)}}$ (SD)	$\hat{T}_{0.05}$ (95% CI)	$E[N(t)]\|_{t=\hat{T}_{0.05}}$ (SD)
1	492 (4.02)	2636 (12.3)	495 (493, 497)	2654 (13.0)
1.5	490 (4.02)	2639 (12.4)	493 (491, 495)	2657 (13.1)
2	487 (4.02)	2636 (12.3)	490 (488, 493)	2654 (13.0)

Table 8: Predicted closure date and estimated total final accrual at t_1=26APR2006 for varying c in C-08

c	$\widehat{E(T)}(SE)$	$E[N(t)]\|_{t=\widehat{E(T)}}$ (SD)	$\hat{T}_{0.05}$ (95% CI)	$E[N(t)]\|_{t=\hat{T}_{0.05}}$ (SD)
1	493 (4.10)	2640 (12.5)	495 (493, 498)	2652 (12.9)
1.5	490 (4.10)	2637 (12.4)	493 (491, 496)	2655 (13.1)
2	488 (4.10)	2640 (12.5)	490 (488, 493)	2652 (12.9)

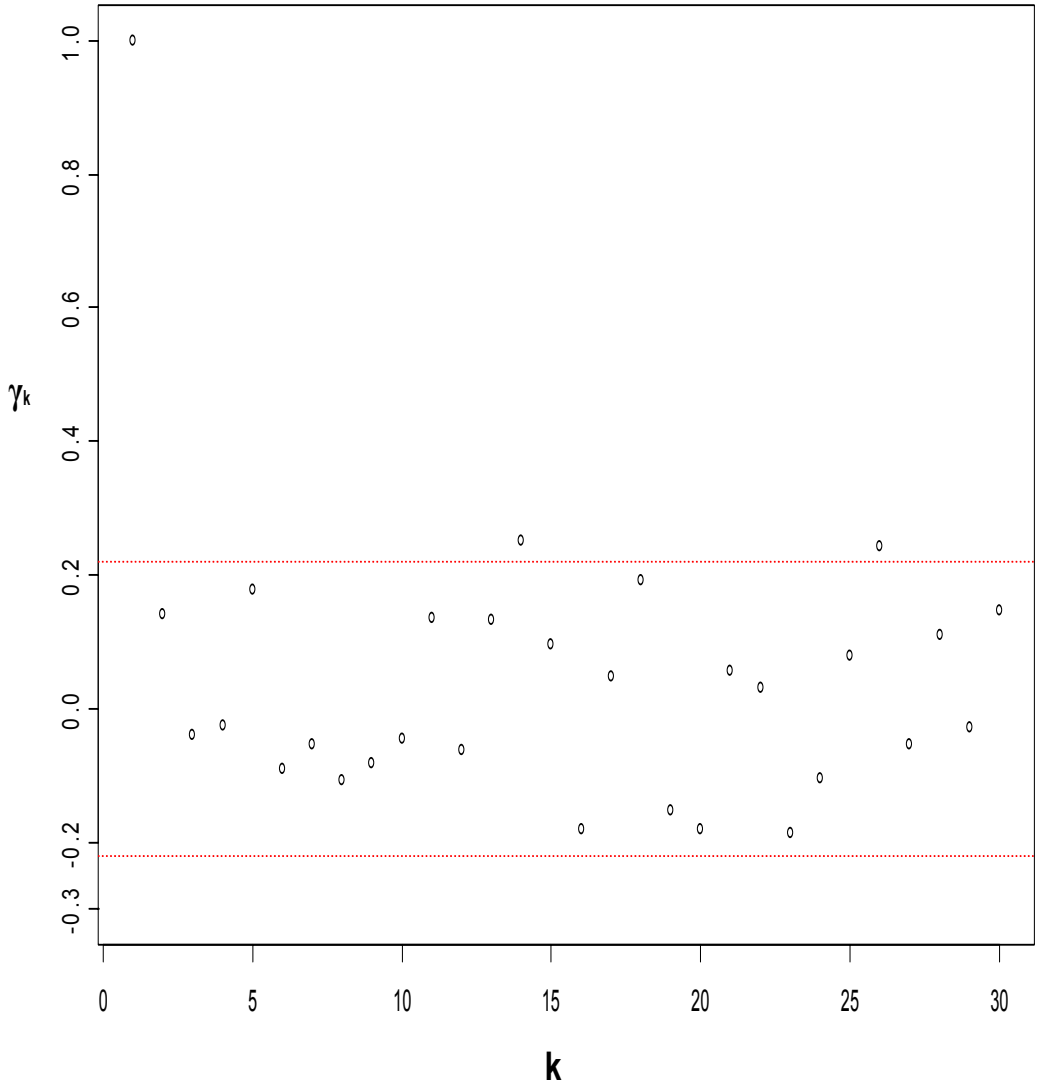

Figure 14: Overall correlogram for C-08, t_1=26APR2006.

5.0 CHANGE-POINT ANALYSIS

5.1 INTRODUCTION

As seen in previous chapters, the choice of the cut-off point t_1, where the daily accrual rate became stable and extended to current time t_2, is critical to the estimation of average accrual rate and determination of the trial closure date. Given t_1, the sample mean should be close to the sample variance for the accruals during $[t_1 + 1, t_2]$ if the accruals during that period are generated from a homogeneous Poisson process. Therefore whether the ratio between the sample mean and the sample variance, for accruals during a period, is close to 1 or not can be used to guide the choice of t_1. Furthermore the goodness-of-fit test statistic G^2 is also informative in finding appropriate t_1. These tools were implemented in the previous chapter. However, they are rather heuristic approaches. Here we are applying change point analysis to find the cut-off t_1 when the accrual rate would become stable and homogeneous up to the current time t_2.

Change point analysis is to find one or more locations where a parameter of a statistical model changes its value. Such analyses have been applied in analysis of data collected from the fields of finance, econometrics, software development, and medicine (Page, 1955 [14]; Chernoff and Zacks, 1964 [7]; Bhattacharyya and Johnson, 1968 [4]; Goldfeld and Quandt, 1973 [10]; Hinkley, 1970 [12]; Zhao and Wang, 2007 [17]). Akman and Raftery (1986) [2] studied the change point problem in a continuous time Poisson process. Hinkley (1970) [12] discussed the maximum likelihood method for estimating a change point in a sequence of random variables, where the probability distribution changes, and the likelihood ratio test statistic. Testing the existence of one or more change points in the means of a sequence independent Poisson-distributed random variables was investigated by Henderson and Matthews

(1993) [11] and the general approaches for change point problems were discussed in Bhattacharya (1994) [5].

Regarding the daily accruals from a clinical trial as a sequence of Poisson-distributed independent random variables, we will apply the maximum likelihood method to find the change point where the means of these variables change.

5.2 LOCATING THE CHANGE POINT FOR A SEQUENCE OF POISSON-DISTRIBUTED RANDOM VARIABLES

Consider a sequence of independent Poisson random variables $\{y_k, k = 1, 2, \ldots, n\}$ with means $E[y_k] = \mu_k$. If these y_k's are generated from a homogeneous Poisson process, then μ_k's are equivalent. However, if the corresponding Poisson process is non-homogeneous, μ_ks would change with k. Assume that μ_ks only change once, that is, $\mu_1 = \mu_2 = \ldots = \mu_\tau \neq \mu_{\tau+1} = \ldots = \mu_n$, where τ is the change point. The likelihood function is

$$\begin{aligned} L(\mu_1, \mu_2, \tau) &= \prod_{i=1}^{\tau} f(y_i; \mu_1) \prod_{i=\tau+1}^{n} f(y_i; \mu_2) \\ &= \prod_{i=1}^{\tau} \frac{e^{-\mu_1} \mu_1^{y_i}}{y_i!} \prod_{i=\tau+1}^{n} \frac{e^{-\mu_2} \mu_2^{y_i}}{y_i!} \end{aligned}$$

With τ fixed, the MLEs for (μ_1, μ_2) are $(\widehat{\mu}_{1\tau}, \widehat{\mu}_{2\tau}) = (\sum_{i=1}^{\tau} y_i/\tau, \sum_{i=\tau+1}^{n} y_i/(n-\tau))$. Therefore the logarithm of the profile likelihood of τ is

$$l_{pl}(\tau) \propto \{log(\sum_{i=1}^{\tau} y_i) - log\tau\} \sum_{i=1}^{\tau} y_i + \{log(\sum_{i=\tau+1}^{n} y_i) - log(n-\tau)\} \sum_{i=\tau+1}^{n} y_i. \tag{5.1}$$

The MLE of τ is $\widehat{\tau} = arg\ max_{1 \leq \tau \leq n-1} l_{pl}(\tau)$. One of the regularity conditions for the maximum likelihood method for data with change point is that τ/n converges to a constant, as $n \to \infty$. Under the circumstance that the likelihood function $l_{pl}(\tau)$ has multiple modes, the MLE is often chosen as the mode that is not close to either the end of the period. The derivation of a confidence set for $\widehat{\tau}$ is very complicated (Bhattacharya, 1994) [5].

5.3 TEST FOR A CHANGE POINT

For a given $\tau \in \{1, 2, ..., n-1\}$, it is of interest to test whether τ is truly a change point. Two test statistics are often considered for this purpose. The first is the likelihood ratio test for the following hypothesis test:

$$H_0 : \mu_1 = \mu_2 = \ldots = \mu_n \;\; against \;\; H_A : \mu_1 = \mu_2 = \ldots = \mu_\tau \neq \mu_{\tau+1} = \ldots = \mu_n.$$

Let

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}, \quad \bar{y}_r = \frac{\sum_{i=1}^{r} y_i}{r}, \quad \bar{y}'_{n-r} = \frac{\sum_{i=r+1}^{n} y_i}{n}.$$

The likelihood ratio test for the hypothesis of no change (H_0) against the alternative of changing after τ is:

$$\begin{aligned} LRT &= \{\prod_{i=1}^{\tau} \frac{e^{-\bar{y}_\tau} \bar{y}_\tau^{y_i}}{y_i!} \prod_{i=\tau+1}^{n} \frac{e^{-\bar{y}'_{n-\tau}} \bar{y}'^{y_i}_{n-\tau}}{y_i!}\} / \prod_{i=1}^{n} \frac{e^{-\bar{y}} \bar{y}^{y_i}}{y_i!} \\ &= \bar{y}_\tau^{\tau \bar{y}_\tau} \bar{y}'^{(n-\tau)\bar{y}'_{n-\tau}}_{n-\tau} / \bar{y}^{n\bar{y}}. \end{aligned} \tag{5.2}$$

When $n \to \infty$, $2log(LRT)$ follows the χ^2 distribution under the null.

The other useful test statistic is the following:

$$T = \frac{\bar{y}_\tau - \bar{y}'_{n-\tau}}{\sqrt{\bar{y}_\tau/\tau + \bar{y}'_{n-\tau}/(n-\tau)}}. \tag{5.3}$$

When $n \to \infty$, T is asymptotically $N(0, 1)$ under the null.

5.4 APPLY THE CHANGE-POINT ANALYSIS IN B-38 AND C-08

In the NSABP trial B-38, the pre-determined target sample size was 4800. Since the required sample size is relatively large, patients were accrued from hundreds of institutions and sites over time. During a work day, there is a small chance for a single institution to put a patient on trial. Therefore, patient accrual during a certain time period is mostly independent from the accrual during another disjoint time period and it approximately follows a Poisson process. Denote the daily accrual over time k by y_k and daily accrual rate by $E[y_k] = \mu_k, \quad k = t_0, t_0 + 1, \ldots, t_2$, where t_2 is the time when we need to make the prediction of the future closure date based on the observed accrual pattern from t_0 up to t_2.

In the proposed methods for predicting the final closure date, the accrual process during $[t_1, t_2]$ is assumed to be a homogeneous Poisson process and observed accrual daily counts y_ks are a sequence of independent Poisson variables with a common mean. Then the observed daily counts during $[t_1, t_2]$ form the basis of prediction. Therefore an appropriate choice of t_1 is critical.

At the beginning of patient accrual, the daily accrual tended to be slow because many participating institutions needed time to gear up their effort on patient accrual. The accrual started getting stable after March of 2005. Our strategy is to find an initial date when the daily accrual was stabilized, then find out whether there was a change point in daily accrual rate with regarding the accrual counts, after that initial date and prior to t_2, as a sequence of independent Poisson random variables.

For the change-point analysis on B-38, we pick the starting point on July 15, 2005, based on the accrual pattern from Figure 1. The potential change point τ will be between t_{175} (15July2005) and t_{593} (19Mar2007).

Figure 15 shows the logarithm of the profile likelihood of τ over time. On December 29, 2006, it has the largest value. The change-point analysis reaches the same result as the algorithm we proposed in Section 4.1.2 based on the ratio between the sample mean and the sample variance and the goodness-of-fit test statistic G^2.

After identify τ, we could also test whether τ is truly a change point using the likelihood ratio test 5.2 or the test statistic 5.3. The null hypothesis is $H_0 : \mu_1 = \ldots = \mu_n$, against the

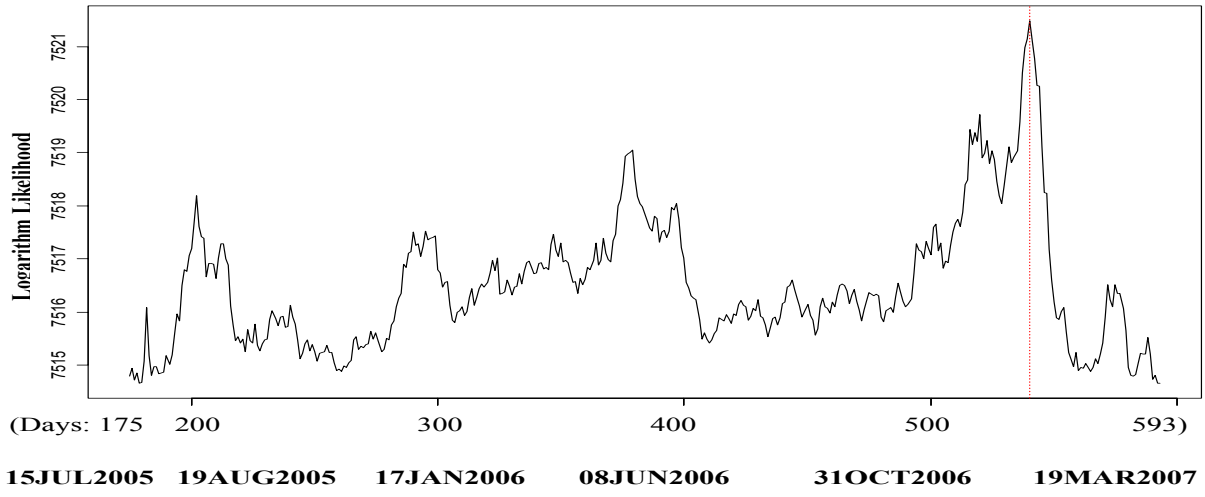

Figure 15: The logarithm of the profile likelihood over time for NSABP trial B-38.

alternative hypothesis: $H_\alpha : \mu_1 = \mu_2 = \ldots = \mu_\tau \neq \mu_{\tau+1} = \ldots = \mu_n$ for some $1 \leq \tau \leq n$. Given τ=540 (29Dec2006), the 2log(LRT)=13.66 with p-value$<$ 0.001 and the other test statistics T=-7.83 (p-value$<$ 0.001). Therefore, $\widehat{\tau}$=540 (29Dec2006) is a change point for accruals from t_{175} (15July2005) to t_{593} (19Mar2007).

For NSABP trial C-08, the pre-determined target sample size was 2632. For the change-point analysis on accrual data from C-08, we picked the starting point as May 19, 2005, based on the accrual pattern from Figure 8. The potential change point τ will be somewhere from $t = 150$ (19May2005) to $t = 467$ (23Aug2006).

Figure 16 shows the logarithm of the profile likelihood of τ over time. There are multiple modes observed. The global mode was only a couple of days ahead of t_2. We did not pick t_1 at that time. As stated earlier, this is required by the large sample property of MLE for change point problems. We picked the first relatively large mode of the logarithm likelihood is on May 26, 2006 ($\widehat{\tau} = 406$). The change-point analysis did not reach the same result as the procedure which we proposed in Section 4.1.2 based on the ratio between the sample

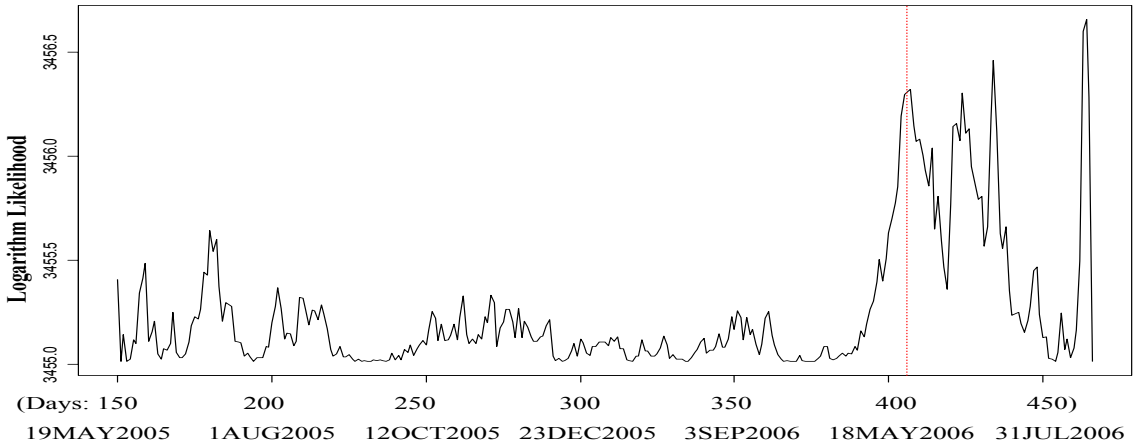

Figure 16: The logarithm of the profile likelihood over time for NSABP trial C-38.

mean and the sample variance, and the goodness-of-fit test statistic G^2. This is probably due to the volatile accrual of C-08 and one change point is not adequate for finding when the accrual rate started being stabilized at the same rate at t_2.

After identify τ, we could also test whether τ is truly a change point using the likelihood ratio test 5.2 or the test statistic 5.3. The null hypothesis is $H_0 : \mu_1 = \ldots = \mu_n$, against the alternative hypothesis: $H_\alpha : \mu_1 = \mu_2 = \ldots = \mu_\tau \neq \mu_{\tau+1} = \ldots = \mu_n$ for some $1 \leq \tau \leq n$. Given $\widehat{\tau}$=406 (26May2006), 2log(LRT)=2.59 with p-value 0.068, and the other test statistics T=4.83 (p-value$<$ 0.001). At t_{406}, the ratio of sample variance to mean is 1.06, which is relatively close to 1. Moreover, the p-value of G^2 for the given t_1 is 0.288, which is not significant and relatively large. Even though the LRT does not suggest that May 26, 2006, is a real change point, these results still indicates that it is an appropriate choice of t_1. Also, given t_1 is chosen on May 26, 2006, the predicted closure date and estimated total final accrual, in Table 9, comparing to the results given t_1 on April 26, 2006 (Table 8) are very similar. For the projected closure date, it takes couple additional days for the sites to

Table 9: Predicted closure date and estimated total final accrual at t_1=26MAY2006 for varying c in C-08

c	$\widehat{E(T)}(SE)$	$E[N(t)]\|_{t=\widehat{E(T)}}$ (SD)	$\hat{T}_{0.05}$ (95% CI)	$E[N(t)]\|_{t=\hat{T}_{0.05}}$ (SD)
1	494 (4.70)	2635 (12.3)	497 (495, 501)	2652 (13.0)
1.5	492 (4.70)	2638 (12.4)	495 (492, 498)	2655 (13.1)
2	489 (4.70)	2635 (12.3)	492 (490, 496)	2652 (13.0)

accrue the required target sample size given the same value of c when t_1 is chosen on May 26, 2006. However, the predicted total accruals are almost the same. It indicates when there are multiple modes of the logarithm of the profile likelihood of τ over time and they are relatively close to each other, it is safe to choose the earliest time point for t_1 which ensure the large sample property of MLE for change point problems.

6.0 EXTENSION

6.1 INCORPORATE SITE INFORMATION OR COVARIATES

If for each participating site, the time when an IRB approval is obtained is available, the Animisov and Fedorov (2007) [3] approach may be modified to get more precise estimate of μ. However, the difference resulted in prediction of final closure date would be minimal if there is any.

In usual multi-institutional clinical trials, not only daily accrual $x(t)$'s are available but also some demographic factors Z such as age and other prognostic factors are also collected. A more general Poisson process model is to incorporate such information into account by modeling $\lambda(t \mid z)$:

$$\lambda(t \mid z) = \lambda_0(t) + \beta(t; z),$$

where care is needed for identifiability. Such models can be used to monitor the accrual rates for various subsets of accrual population continuously. The results can be used as basis for whether action should be taken to improve the accrual of specific subsets. It can provide a better idea about the accrual pattern of sub-populations. Also, it can lead to more precise estimate of closure date because more data are used.

6.2 WHEN SERIAL CORRELATION EXISTS

When there are serial correlations among the accrual, a reasonable extension is to assume that

$$X(t) \sim Poisson(\mu + \delta(t)),$$

where $\delta(t)$ is a Gaussian process with correlation:

$$corr(\delta(t), \delta(s)) = exp(-\phi \mid t - s \mid).$$

However, the likelihood function of μ does not have an analytical form and numerical integration is required. The prediction of future closure date would be very complicated, because the joint distribution of the future daily accruals would be complex. Then a Bayesian method would be more suitable for the purpose of predicting the future closure date.

APPENDIX

THE CORRESPONDING ACTUAL ACCRUALS

Table 10: Dates close to the trial closure and the corresponding actual accruals in B-38

Parameters	Date	Days	Accumulative Accrual
t_0	November 2nd, 2004	0	0
t_2	March 20th, 2007	594	4465
Potential t_3	April 18th, 2007	615	4793 *
	April 19th, 2007	616	4802 *
	April 20th, 2007	617	4811 *
	April 23th, 2007	618	4819 *
	April 24th, 2007	619	4832 *
	April 25th, 2007	620	4842 *
	April 26th, 2007	621	4849 *
	April 27th, 2007	622	4860 *
	April 30th, 2007	623	4872 *
	May 1st, 2007	624	4881 *
	May 2nd, 2007	625	4888 *
	May 3rd, 2007	626	4894 *

* These accumulative accruals were calculated as the observed total accrual during the last five days prior to May 3, 2007, and the period between $t_2 + 1$ and $t_3 - 5$.

Table 11: Dates close to the trial closure and the corresponding actual accruals in C-08

Parameters	Date	Days	Accumulative Accrual
t_0	September 15th, 2004	0	0
t_2	August 23th, 2006	467	2489
Potential t_3	September 20th, 2006	486	2639 *
	September 21st, 2006	487	2647 *
	September 22nd, 2006	488	2653 *
	September 25th, 2006	489	2664 *
	September 26th, 2006	490	2669 *
	September 27th, 2006	491	2677 *
	September 28th, 2006	492	2681 *
	September 29th, 2006	493	2684 *
	October 2nd, 2006	494	2694 *
	October 3rd, 2006	495	2697 *
	October 4th, 2006	496	2700 *
	October 5th, 2006	497	2707 *
	October 6th, 2006	498	2710 *

* These accumulative accruals were calculated as the observed total accrual during the last five days prior to October 6, 2007, and the period between $t_2 + 1$ and $t_3 - 5$.

BIBLIOGRAPHY

[1] Agresti, A. Categorical data analysis. *Wiley Interscience*, **2nd ed.**,2002.

[2] Akman VE and Raftery AE, 1986. Asymptotic inference for a change-point Poisson process. *Ann. Statist.*, **14**, 1583-1590.

[3] Anisimove, VV and Fedorov, VV (2007). Modeling, prediction and adaptive adjustment of recruitment in multicentre trials, *Statistics in Medicine*, **26**, 4958–4975.

[4] Bhattacharyya, GK and Johnson, RA (1968). Non-parametric tests for shift at an unknown time point. *Ann. Math. Statist.*, **39**, 1731-1743.

[5] Bhattacharyya, PK. Some Aspects of Change-Point Analysis. *Change-point problems: Papers from the joint AMS-IMS-SIAM Summer Research Conference. South Hadley, MA: 1992.* Editor: E. Carlstein, H. Mller, and D. Siegmund, (Lecture Notes–Monograph Series), **23**, 1994. Hayward, CA: Institute of Mathematical Statistics.

[6] Cook, RD and Weisberg, S. Residuals and influence in regression. *Chapman and Hall: New York and London*, 1982.

[7] Chernoff, H and Zachs, S (1964). Estimating the current means of a normal distribution which is subjected to changes in time. *Ann. Math. Statist.*, **35**, 999-1018.

[8] Diggle, PJ. Time series: A biostatistical introduction. *Oxford University Press: New York*, 1990.

[9] Gajewski, BJ, Simon, SD, and Carlson, SE (2008). Predicting accrual in clinical trials with Bayesian posterior predictive distributions, *Statistics in Medicine*, **27**, 2328-2340.

[10] Goldfeld, SM and Quandt, RE (1973). A Markov model for switching regressions. *Journal of Econometrics*, **1**, 3-16.

[11] Henderson R and Matthews JNS. (1993) An Investigation of Changepoints in the Annual Number of Cases of Haemolytic Uraemic Syndrome. *Applied Statistics*, **42**, 461-471.

[12] Hinkley, DV (1970). Inference about the change-point in a sequence of random variables. *Biometrika*, **57**, 1-16.

[13] Kingman, JFC. Poisson processes. *Oxford University Press: New York*, 1993.

[14] Page, ES (1955). A test for a change in a parameter occurring at an unknown point. *Biometrika*, **42**, 523-526.

[15] Senn, S (1998). Some controversies in planning and analyzing multi-centre trials. *Statistics in Medicine*, **17**, 1753-1765.

[16] Taylor, HM and Karlin, S. An introduction to stochastic modeling. *Academic Press*, **3rd ed.**,1998.

[17] Zhao, J. and Wang, J. (2007) Testing the existence of change-point in NHPP software reliability models. *Communications in Statistics-simulation and Computation*, 36, 607-619.