UNIVERSITY OF NAIROBI

SCHOOL OF MATHEMATICS

USING SEQUENTIAL ANALYSIS TO SHORTEN CLINICAL TRIALS

By

Okemwa Benard Onserio

156/61414/2013

A research project submitted in the partial fulfilment of the requirement for the award of the degree of Master of Science in Biometry

JULY 2014

DECLARATION

This project is my original work and has not been submitted for a degree or any other academic institution of Higher Learning.

Signature……..Date……

Okemwa Benard Onserio.

I56/61414/2013

Approval

This is to certify that the project titled "Using sequential Analysis To shorten Clinical Trials" carried out by the above named student has been read and approved for meeting part of the requirements governing the award of Masters of Science in Statistics(Biometry) degree of the university of Nairobi.

Supervisors:

Signature……..Date……

Dr. Nelson Onyango Owuor

Signature……..Date……

Mrs. .Anne Wangombe

DEDICATION

This project is dedicated to the Ministry of Labour and National Industrial Training Authority for sponsoring my Masters of Science Education, my wife and my children inspirations and self collectively have made my dream a reality

ACKNOWLEDGEMENT

I would like to acknowledge with gratefulness the endeavour from my supervisors Dr. Nelson Onyango Owuor and Mrs. Anne Wangombe for their guidance, patience, encouragement and overwhelming support provided to me throughout this work.

I am profoundly grateful to entire University of Nairobi, in particular Chiromo Campus –School of Mathematics staff for their support at whatever levels by making completion a reality.

Finally, it indeed my great pleasure to appreciate the support to me by my fellow students during the study period.

Above all I thank the Almighty God for the perfect care throughout the period.

TABLE OF CONTENTS

ACRONYNAMES

CSP - cyclosporine.

GVHD - Graft Versus Host Disease.

MTX - Metroxate.

ASN - Average Sample Number

ABSTRACT

Sequential analysis is concerned with the statistical theory and methods of analysis of such data in which the final number of observations need not to be fixed in advance but may depend in some specified way on the data themselves as they become available. One way to sufficiently utilise sequential analysis is using triangular test approach to shorten clinical trials. A situation of accumulation allows analysis of the data as they accumulate and stop the trial as soon as superiority of one or other treatment is established.

The statistical analysis should meet the following ethical requirements

i. If a larger treatment difference becomes apparent during the course of the study, then the study will stop.

ii. If no treatment difference is observed, the trial will continue to its maximum sample size.

iii. The assumptions of the study is that the new treatment will be found to be more efficacious compared to control (standard) treatment. A bio equivalence study gives the hope that the new treatment will have equivalent therapeutic value such that its advantage may lie elsewhere perhaps it is cheaper to produce, easy to administer or fewer unpleasant side effects.

CHAPTER ONE - INTRODUCTION

1.1 Background of the Study

A clinical research are medical experiments which involves human beings, they need accurate understanding of the statistical design that can give valid analysis for decision making. Ethical standards are of considerable importance to be considered. A study which may withhold prove of inferiority in the earliest may be stopped and not to progress to definitive stage (conclusion).

Sequential analysis offer an opportunity to stop a trial as soon as the evidence for or against the treatment has reached the conventional levels required for the scientific study. The random experiment usually yields observations (data) say X_1 $X_{2...}$ that are random variables (or vectors) not necessarily independent to be analyzed. The end of the experiment depends on:

1. Stopping rule, which dictates whether experimentation should be stopped with $(X_{1,}\ldots,X_n)$ or continued with the additional observations $(x_{n+1)}$ for each $n \geq 1$and
2. A decision rule that tells what terminal action is to be taken about the given problem after experimentation has stopped.

A stopping rule gives rise to a stopping variable "N" which gives an extended random variable such that for n =1, 2,…,∞, the event N= n is the sigma fields (ϵn)generated by ($x_1 \ldots x_n$)

A decision rule D is then an εN measurable function taking values in some well-defined space.

The goal of a sequential analysis is then to determine an optimum (N, D) or under a specified N and "optimum" D that meets a certain desirable criteria.

1.2 Problem Statement

Taking full time schedule to test efficacy of a drug before it is availed asstandard treatment of people (human) denies the patients the opportunity to use it early unnecessarily if interim results can proof sustainable efficacy.

1.3 Justification of the Study

Saving lives is a central interest for any scientific idea and clinical trial which attempt to address existing and out breaking medical conditions.However it takes relatively and unnecessarily long time before a new, superior and alternative treatment is availed for human .It is for this reason that sequential statistical approach can fully address.

1.4 Main Objectives of the Study

1. To use sequential analysis on interim results to shorten a clinical trials. This means analyzing expected results as they are available and if they satisfying statistical authorities the decision can be made early
2. To stop trials and uphold the new drug as superior or has satisfied that: if consequent trials demonstrate negligible or no additional information or significant difference in the earliest met expectation of the drug.

CHAPTER TWO - LITERATURE REVIEW

2.0 Introduction

Sequential design was derived from walds work on sequential probability ratio by Armitage (1960) significant tests were limited by the need to update sequential analysis after every patient response and by the need to recruit patients in pairs for one to receive the experimental treatment and the other the control.

Choenfeld (1980) suggested a conventional hypothesis testing formulation in which the test therapy is compared with a standard treatment.Large samples necessary to convince the concerned .conventional designs are described by peacock (1983) who takes the standard hypothesis testing approach. The parameter of interest is a measure of difference inefficiency between the two treatments on suitable scale.

2.1 Hypothesis

Null hypothesis (H_0)$= no$ difference

Alternative hypothesis (H_a) $=$ there is difference.

For specified degree of therapeutic improvement the power of the test to achieve a stated significance level is fixed and the sample size is determined.

It involves accumulation of subjects gradually over a period of time .the results from early subjects are available at a time when subsequent patient are being recruited. Such a situationaccumulates and stops the trial as soon as superiority of one or the other treatment is established. Kilpatrick and Oldham (1954) used a t- test in comparison of two bronchial dilators successfully and that the stopping rule was so successful that the trial was terminated after only four patients' sequential plan.

The above constrain was sorted out by group sequential analysis by Pocock (1977,1982) and O Brien and Fleming (1979) where Pocock, O Brien and Fleming proved that the future of the trial is determined at a limited number of inspections (looks) at the data after each new group of subjects have responded.

McPherson (1982) discusses the optimum number of looks to make. The group sequential designs of pocock are essentially repeated significance tests, but the individual tests are applied after a specified group of patients have responded and not after every patient. Designs of this form have been used in large scale investigations concerning life threatingdiseases inorder to satisfy safety monitoring needs without compromising the eventual statistical analysis. The ethical requirements of the design are

i. If a large treatment difference becomes apparent during the course of the study then the studywill stop.

ii. If no treatment difference is observed the trial will continue to its maximum sample size.

The assumption was that the new treatment will be found to be more efficacious compared to the control (standard treatment) in case of a bio equivalence study the hope is that the new treatment will have equivalenttherapeutic value. If so, its advantage may lie elsewhere perhaps it is cheaper to produce, easy to administer or fewer unpleasant side effects.

2.2 A Study of Immune Suppressive In Bone Marrow Transplantation.

Is a case study of clinical trial reported by Stub et al., (1986) it was conducted at the Fred Hutchinson cancer research Centre in Seattle, Washington. The Centre transplanted bone marrow to blood cancer patients. The trial in question was a comparison of two immuno- suppressive treatments used prior to transplantation in order to reduce the likelihood of graft rejection.

Response of interest is failure time (survival analysis) assuming proportional hazards model. Suppose that patients on the new treatment (CSP +MTX) have a hazard function $h_N(t)$ and survival function $F_N(t)$, while those on the standard treatment (CSP alone) have Hazard function $h_S(t)$ and survival

function $F_S(t)$. the measure of treatment difference will be taken to be θ Where

$$h_N(t) = \exp(-\theta)\, h_S(t), \text{ For all } t > 0$$

The treatment difference θ can be given in terms of survival function

$$\theta = \log\{-\log F_S(t)\} - \log\{-\log F_N(t)\}$$

And the sequential monitoring procedure involves a plot of the log rank statistics 'Z' against its null Variance 'V'.

Where:

$$Z = Ms - \sum_{i=1}^{k} o_i r_{is} / r_i$$

$$v = \sum_{i=1}^{k} \frac{o_i(r_i - o_i) r_{is} r_{iN}}{r_i^2 (r_i - 1)}$$

2.2.1 The Stopping Boundaries

Using a triangular test by Whitehead (1983), the general form of the continuation region or the final triangular test is given by

$$Z\epsilon(-a + \gamma v\,, a + \mu v)$$

Where $\gamma = 3\theta_R/4, \mu = \theta_R/4,$

$$a = \left(\frac{2}{\theta_R}\right) \log \left(\frac{1}{2\alpha}\right)$$

From the observations plotted:

$$Z = -a + v, \rightarrow for\ Z = -6.1571 + 0.561v \text{ (Lower boundary)}$$

and

$$Z = 6.15 + 0.187v \text{ (Upper boundary)}$$

Alternatively he also used Flexible monitoring; because it is not possible to monitor the progress of atrial continuously a flexible monitoring plan was developed by Whitehead (1983). This was the explanation of the plan:

Denote the value of statistics Z and V collected at the i^{th} inspection by Z_i and V_i and let$V_0 = 0$, at the i^{th} inspection the test will terminate with the conclusion that the new is superior if:

$$Z_i + 0.583\sqrt{v_i} - v_{i-1} > 6.15 + 0.187v_i,$$

And will terminate and accept the null hypothesis if:

$$Z_i - 0.583\sqrt{v_i} - v_{i-1} < -6.15 + 0.562v_i$$

Otherwise it will continue.

The derivation of the constant 0.583 is described by Saike and Sigmund (1985) flexible monitoring preserves the significance level and power requirements or the trial to a remarkable degree of accuracy.

Terminal analysis: Because the upper boundary had been crossed it was apparent that CSP+MTX were significantly superior to CSP alone at the 5% significance level one sided alternative.

The study terminated after the eight inspections when 79 patients had been recruited and 49 of these had failed. This implied a saving in sample size and shortened a clinical trial.

2.3 Repeated Significance Testing and Early Rejection of H_0

Group sequential analysis is done on the basis of repeatedly performing fixed sample analysis in an organized way. The statistical approach requires:

i. Defining a formal stopping rule in terms of first kind error risk Test power
ii. Considering its relationship with the number of looks and the total (maximum) size of the trial

Considered a homogeneous sample of patients sequentially entering into a clinical trial aiming at comparing the effects of the two treatment A and B. by a randomized permuted block design each consecutive set "2n" patient is

assigned to the treatments so that n of them are allocated to A and B respectively.

Assume the response a Gaussian random variable with a known variance and unknown average $\mu_A \; and \; \mu_B$ and that $\pi_{Aj} \; and \; \pi_{Bj}$are corresponding observed mean responses in the J^{th} set of n subject, then theempirical difference computed at the i^{th} look is represented by :

$$dj = \sum_{j=1}^{j} (\pi_{Aj} - \pi_{Bj})$$

Where i= 1, 2…L

and L= number of anticipated looks for a given trial and its Gaussian with expectation.

$$id = i(\mu_A - \mu_B) and$$

Variance $\frac{2i\sigma^2}{n}$

From Pocock (1977), a two- tailed significance test of the null hypothesis.

$$H_0 \colon \; \delta = 0$$

Computed at the i^{th} look has probability.

$$p1 = 2[1 - \emptyset\{\sqrt{nd(i)} \;\; / \sigma\sqrt{2i}) \}$$

When standardized Gaussian distribution repeated testing of each look increases the probability of finding a statistically significant difference purely by chance (Armitage et al., 1969) is realized, it is assumed that after achieving a given nominal level of significance, the investigator stops the study.

2.3.1 Different Strategies Aiming At Making the Final Close to by Applying Benferronis Inequality

One may control the probability ofα by rejecting the null hypothesis on the whole set of test when in fact it is time, by adopting more stringent nominal significant levels α for the test compute at each look. According to Pocock (1977) givesα_1^1 values and their corresponding Gaussian deviates Z (α, L) for looks between 1 + L = 20. They have been determined so that the overall risk of first kind error is:

$$\alpha = 0.05 \; or \; \alpha = 0.01$$

Note that given α and L.α^1 values do not depend on the number n of subjects whose responses are observed at each look because the differences ($\pi_{Aj} - x_{Bj}$) are random variables identically and normally distributed with variance $2\frac{\sigma^2}{n}$.

Generally the two tailed group sequential procedure is:

a) Reject H_0 and stop the trial at the i^{th} look if$(\frac{d(i)}{\sigma}) > Z(\alpha, l)\sqrt{2i/n}$. otherwise continue the trial .comparison of three group sequential designs in terms of average sample number (ASN) Pocock's design is assumed as a reference design, (two tailed) tests with and five looks.

b) Accept H_0 if at last planned look

$$[\frac{d(L)}{\alpha}] \leq Z(\alpha, L)\sqrt{\frac{2L}{n}}$$

2.3.2 Interpretation

In interpreting the results of a clinical trial in which repeated significance testing at a constant nominal level α has been adopted, adilemma may arise when none of the consecutive tests reaches α but the final "P" value is less than the overall significance level α^1. In such a case suggestions aiming at making final α_1^1 close to α could be given. For instance after choosing a pertinent value of α_1^1 close to one could determine the remaining value of α^1 using Benferronis inequality.

The most relevant feature of a sequential design is known to which it allows the trial to stop early when the alternative hypothesis is true from this view point the most informative statistical quantity is the mean number of observations required to reach a decision.

2.4 A Symmetric Group Sequential Boundaries and Early Accept of H_0

Demmets and Were (1985) have shown that in one sided settings one can take advantage of the stopping rules which enable early termination and acceptance of H_0 at each look. In case we consider $H_0 : \sigma \leq 0 \; and \; H_a \; : \sigma > 0$ and it is sensible to assume that acceptance of Ho may be suggested before the statistics reaches a value as extreme as one needed to verdict Ho hence a symmetric group sequential boundaries are obtained at each look.

2.5 Ian and Demmets Procedure

Ian and Demets (1983) generalizes the group sequential approach allowing the investigator to escape the two restrictions i.e. specifying the number of interim analysis and performing analysis on accrual of information an but resorted to properly built boundaries.

This procedure relies on "α spending function"i.e.let's assume that the trial be completed by time T, stated arbitrary such that the T $=1$ .The$\propto (t)$ function allocates the amount of type 1 error that the investigator can spent at each analysis carried out at the $t(o \leq t \leq 1 \leq$ and is such that $\alpha(o) = 0$ and $\alpha_1 = \alpha$.

Parameter t defines the position of the interim analysis during the trial and if n_1 patients have accrued before the i^{th} analysis,

$$t_i = \frac{n_i}{N} (where\ N\ is\ the\ target\ sample\ size)$$

Thus the final analysis corresponds to T =L. if by construction the increment $\alpha(t_1) - \alpha(t_{1-1})$ is the amount of significance level the investigator can utilize at the time t_1 for the whole trial the risk of type 1 error is α regardless of the data or the number of interim analysis .The evaluation of boundary values requires numerical integration $b_i = Z_i\sqrt{t_i}$ which can give early stoppage.

2.5.1 Stochastic Curtailed Sampling

The procedures considered so far base the recommendation to stop the trial on the current evidence about the size of the relevant parameter in relation to properly defined critical values. Scholastic curtailment (Ian et al.,1982) takes into account information collected at a given interim analysis, it attempts to predict the final results that would be obtained if the trial were allowed to complete its course until time T .clearly the future data are unknown and the final results are subject to random uncertainty. However, if at the i^{th}look it becomes known that, with high probability the future data could not change the conclusion could be reached. It is sensible to stop the trial at this point.

2.6 Two sided tests introduction

Two sided tests for comparing two treatments with normal response of known variance. Two-sided test is given to test hypothesis against a two sided alternative .our attention is focused on testing for difference in the mean response of two treatments when observations are normally distributed with common known variance. Denoting the difference means by θ the null hypothesis.

$H_0: \theta = 0 \rightarrow$ States those responses follow the same distribution under both treatments.

$H_a: \theta \neq 0$ Contains two cases which correspond to one treatment being superior to the other and vice versa

In this comparison, the standardized test statistic Z, is distributed symmetrically about 0.under H_0 and affixed sample test rejectH_0 if |Z|>C for some constant C. The sign of Z determines which treatment is to be preferred when Ho is rejected. A tests type 1 error probability is defined to be. The probability of wrongly rejecting the null hypothesis:

$$P_r = \theta\{|Z| > C\}$$

The power of test is the probability of rejecting the null hypothesis when it does nothold $P_{r\theta}\{|Z| > C$ for values of $\theta \neq 0$

2.7 Malaria Journal 2002

Triangular test applied to the clinical trial of aAzithromycin against relapses in plasmodium viral infection. Triangular test applied to the clinical trial of a zithromycinagaist relapses in plasmodium virax infections.

Administration: Patients infected with P. virax were treated with a zithromyan 1.2g daily for 7 days after that treatment; the onset of a relapse infection was monitored.

Results: Five acute cases of P. virax infections were included in the study .all the patients were cured except three patients who reported mild adverse effects.

When the third patient relapsed the sample path crossed the upper boundary of the triangular test the trial was stopped.

CHAPTER THREE - METHODOLOGY

3.1 General Set-Up of a Triangular Test.

The triangular test is non-comparative and is adapted by Belizean et.al .the design aims to compare an observed event rate P to an expected event rate Po.

3.1.1 Hypotheses

$$H_{0:}\, P \leq P_0 \text{ and}$$

$$H_a :\ P > Po$$

The log odds ratio will be used to measure the difference between P and P_0 i.e.

$$\theta = log\{\frac{P(1-P0)}{po(1-p)}\}$$

For instance if $p = po$, Then $\theta = 0$

3.1.2 Illustration

Let $p = 12\% = 0.12$(observed rate)

$$po = 12\% = 0.12(expected\ event\ rate)$$

That is $p = po\ then$

$$\theta = \log\{\frac{0.12(1-0.12)}{0.12(1-0.12)}\}$$

$$= \log\left\{\frac{0.12(0.88)}{0.12(0.88)}\right\}$$

$$= \log 1 = 0$$

Case I: Let's take the observed rate be

$$p = 8\% = 0.08$$

And the threshold (expected)$= 12\% = 0.12 = po$

$$\theta = \log\{\frac{0.08(1-0.12)}{0.12(1-0.08)}\}$$

$$= \log\{\frac{0.08 \times 0.88}{0.12 \times 0.92}\}$$

$$= \log\{0.637681\}$$

$$= -0.1954$$

Case two: Let the observed rate be 45% (large) and the threshold (expected) be 12%

$$\theta = \log\{\frac{0.45(1-0.12)}{0.12(1-0.45)}\}$$

$$= \log\{\frac{0.45(0.88)}{0.12(0.55)}\}$$

$$= \log\{\frac{0.396}{0.066}\} = \log 6 = 0.777815$$

When p (responserate) is large enough 45% compared to Po (expected) the log odds ratio difference is 0.77815.this treatment difference is large enough that no further investigation is necessary. At this level the investigator can make a decision that the new treatment is better in performance.

3.2 Deriving the Equations of Determining the Log-Rank Statistic Z and the Null Variance V.

Get the log likelihood of the data we can use the formula

$$(p) = S\ logp + (N - S)\log(1 - p)$$

Where: S – denotes the number of patients who experience a response (success), N- The number of patients included in the study.

The Z and V statistics are calculated as follows. Where

Z =the efficient score

V = the fisher's information for the parameter of interest. In our case is relapse

$$z = \frac{\partial l}{\partial \theta} at(\theta = 0) = S - NPo$$

$$v = \frac{\partial^2 l}{\partial \theta^2} at(\theta = 0) = NPo(1 - Po)$$

This formula for my methodology can be derived as

$$if\ \theta = \log\{\frac{P(1 - Po)}{Po(1 - P)}\}$$

$$let\ e^{\theta} = (P(1 - Po))/(Po(1 - P))$$

$$e\theta po(1 - p) = p(1 - po)$$

$$e\theta Po - e\theta PoP = P - PPo$$

$$e\theta Po = P - PPo + e\theta PoP$$

$$P - PPo + e\theta PoP = e\theta Po$$

$$P(1 - P0 + e\theta Po) = e\theta Po$$

$$P = \frac{e\theta Po}{1 - Po + e\theta Po}$$

$$P = \frac{eoP0}{e\theta P0 - P0 + 1}$$

$$\text{p}= \frac{e\theta p0}{P0(e\theta - 1) + 1}$$

Given $l(P) = S\ logP + (N - S)\log(1 - P)$

Substituting it on the above i.e.

$$p = \frac{e\theta Po}{Po(e\theta - 1) + 1}$$

$$l(p) = S\log\left\{\frac{e\theta Po}{Po(e\theta-1)+1}\right\} + (N - S)\log(1 - P)$$

Let's find (1-p) given p

$$p = \frac{1}{1} - \frac{e\theta Po}{Po(e\theta - 1) + 1}$$

$$= \frac{Po(e\theta - 1) + 1 - e\theta P0}{Po(e\theta - 1) + 1)}$$

$$= \frac{Poe\theta - Po + 1 - e\theta Po}{Poe\theta - Po + 1}$$

$$\frac{1 - Po}{Poe\theta - Po + 1}$$

Putting together the log likelihood in the data equation, i.e.

$$l(p) = Slogp + (N - S)\log(1 - p)$$

$$= Slog\left\{\frac{e\theta Po}{Poe\theta - po + 1}\right\} + (N - S)\log\{\frac{1 - po}{poe\theta - po + 1}\}$$

Then find the first $partial$ derivative with respect to θ

$$\frac{\partial l}{\partial \theta}$$

$$l(p) = Slog\left\{\frac{e^{\theta}Po}{Poe^{\theta} - Po + 1}\right\} + (N - S)\log\{\frac{1 - Po}{Poe^{\theta} - Po + 1}\}$$

Using the formula or differential calculus

If $\gamma = lnx$ as a function and

$$F = x$$

Then $\frac{dy}{dx} = \frac{1}{F} \times \frac{dF}{dx}$

Applying it on the above relationship our "F"

$$F = \frac{e\theta Po}{Poe^{\theta} - Po + 1}$$

$$\rightarrow \frac{\partial F}{\partial \theta} = \frac{1}{F} \times \frac{\partial F}{\partial \theta}$$

$$\frac{Poe^{\theta} - Po + 1}{e^{\theta}Po} \times \frac{\partial F}{\partial \theta}$$

Get $\frac{\partial F}{\partial \theta}$, using quotient rule, if $y = \frac{u}{v}$ where u and v are functions of x then

$$\frac{dy}{dx} = v\frac{du/dx - u\, dv/dx}{v^2}$$

Then $\frac{\partial F}{\partial \theta}$ will be

Where we let $u = e^{\theta}\, Po$

$$\rightarrow \frac{\partial u}{\partial \theta} = \ e^{\theta}\ Po$$

$$v = Poe^{\theta} - Po\ \ + 1$$

$$\frac{\partial v}{\partial \theta} = \ e^{\theta}\ Po$$

Substituting in the formula or quotient rule

$$\frac{\partial F}{\partial \theta} = \frac{\left(Poe^{\theta} - Po + 1\right) e^{\theta} Po - \left(\ e^{\theta} Po\right) e^{\theta} Po}{(Poe^{\theta} - Po + 1)^2}$$

$$= \frac{Poe^{\theta}(Poe^{\theta} - Po + 1)}{(Poe^{\theta} - Po + 1)}$$

$$= \frac{(Poe^{\theta}(Poe^{\theta - Po + 1 - \left(e^{\theta} p_o\right)^2}}{(poe^{\theta} - po + 1)2}$$

$$= \frac{Pe^{\theta} - P2e\theta}{(P_o e^{\theta} - P_o + 1)^2}$$

Setting together the parts of the formula

Where

$$\frac{\partial l(p)}{\partial \theta} = \frac{1}{F} . \frac{\partial F}{\partial \theta}$$

$$= \frac{Poe^{\theta} - Po + 1)}{Poe^{\theta}} \times \frac{Poe(1 - po)}{(Poe^{\theta} - Po + 1)}$$

$$\frac{\partial l}{\partial \theta} at\theta = 0, S = \frac{1 - Po}{Poe^{\theta} - Po + 1}$$

That was part of the function differentiated i.e.

$$l(p) = SlogP + (N - S)log\ 1 - P$$

$$\log p = \frac{\partial log\ p}{\partial F}$$

$$S\ log\ \{\frac{e^{\theta} po}{poe^{\theta} - po + 1\}}$$

$$\frac{\partial log p}{\partial F} = s\{\frac{1 - po}{po - poe^{\theta} + 1}\}$$

Finding the derivative

$$(N - S)\log\{\frac{(1 - Po)}{Poe^{\theta} - Po + 1}\}$$

Using partial derivative approach, assume the function

$$F = \frac{1 - Po}{Poe^{\theta} - Po + 1}$$

$$\frac{1}{F} = \frac{Poe^{\theta} - Po + 1}{1 - Po}$$

$$\frac{\partial F}{\partial \theta} = by\ quotient\ rule$$

$\theta = \frac{u}{v}$ where u and v are the functions of θ then

$$\frac{\partial F}{\partial \theta} = \frac{v \frac{\partial u}{\partial \theta} - u \frac{\partial v}{\partial \theta}}{v^2}$$

$$u = 1 - Po$$

$$\rightarrow \frac{\partial u}{\partial \theta} = 0$$

$$v = Pe^{\theta} - Po + 1$$

$$\frac{\partial v}{\partial \theta} = poe^{\theta}$$

$$\rightarrow \frac{\partial F}{\partial \theta} = \frac{(Po - Po + 1)0 - (1 - Po)Poe\theta}{(Poe^{\theta} - Po + 1)2}$$

$$= \frac{-Poe^{\theta}\,(1 - Po)}{(Poe^{\theta} - Po + 1)2}$$

Applying

$$\frac{\partial F}{\partial \theta} = \frac{1}{F} \times \frac{\partial F}{\partial \theta}$$

$$= \frac{(Poe^{\theta} - Po + 1)}{(1 - Po)} \times \frac{-Poe^{\theta}\,(1 - Po)}{(Poe^{\theta} - Po + 1)2}$$

$$\rightarrow \frac{\partial F}{\partial \theta} = \frac{-Poe^{\theta}}{Poe^{\theta} - Po + 1}$$

Putting together to get

$$\frac{\partial l(p)}{\partial \theta}$$

$$\frac{\partial l(p)}{\partial \theta} = S\left[\frac{1-Po}{Po\, e^{\theta} - Po + 1}\right] + (N-S)\frac{-Poe^{\theta}}{Poe^{\theta} - Po + 1}$$

Substituting $\theta = 0$

$$S\left[\frac{1-Po}{Po.1 - Po + 1}\right] + (N-S)(\frac{-Po.1}{Po.1 - Po + 1})$$

$$= S\left(\frac{1-Po}{1}\right) + (N-S)(-Po)$$

Opening the brackets

$$= S - SPo - NPo + SPo)$$

$$= S - Npo$$

Hence

$$Z = \frac{\partial l(p)}{\partial \theta} \text{ at}$$

$$\theta = S - NP_0$$

$$v = \frac{\partial^2 lp}{\partial \theta}\ \theta = 0 = NP_0(1 - P_0)$$

Given $\frac{\partial l(p)}{\partial \theta} = S\left[\frac{1-Po}{Poe^{\theta}- Po+1}\right] + (N - S)\frac{-Poe^{\theta}}{Poe^{\theta} -Po+1}$

Then

$$\frac{\delta^2 l}{\delta \theta^2}$$

Can be obtained as follows, using Quotient rule, starting by the one part:

$$F_1 = \frac{1 - Po}{Poe^{\theta} - Po + 1}$$

As the function F:

$F = \frac{u}{v}$, u and v are functions of θ

Then $u = 1 - po$

$$\frac{\partial u}{\partial \theta} = 1 - Po = 0$$

$$v = Poe^{\theta} - Po + 1 , \frac{\partial v}{\partial \theta} = Poe\theta$$

$$\frac{\partial v}{\partial \theta} = \frac{v\partial u/\partial \theta - udu/\partial \theta}{v^2}$$

$$= \frac{\left(Poe^{\theta} - Po + 1\right)0 - (1 - Po)Poe^{\theta}}{(Poe^{\theta} - Po + 1)^2}$$

$$= \frac{-(1-Po)Poe^{\theta}}{(Poe^{\theta}-Po+1)^2} = \frac{Poe^{\theta}\,(Po-1)}{(Poe^{\theta}-Po+1)^2}$$

Going for the second part of

$$F_2 = \frac{-Poe^{\theta}}{Poe^{\theta}-Po+1}$$

$F_2 = \frac{u}{v}$ *where u and v are functions of* θ

$$u = -Poe^{\theta}$$

$$\frac{\partial u}{\partial \theta} = -Poe^{\theta}$$

$$v = Poe^{\theta} - Po + 1$$

$$\frac{\partial v}{\partial \theta} = Poe^{\theta}$$

By

$$\frac{\partial F}{\partial \theta} = \frac{v\frac{\partial u}{\partial \theta}}{v^2} - u\frac{\partial u}{\partial \theta}$$

$$= \frac{\left(Poe^{\theta}-Po+1\right)\left(-Po\,e^{\theta}\right) - \left(-Poe^{\theta}\ .Poe^{\theta}\right)}{(Poe^{\theta}-Po+1)^2}$$

$$= \frac{-\left(Poe^{\theta}\right)^2 - P^2oe^{\theta} - Poe^{\theta} + \left(Poe^{\theta}\right)^2}{(Poe^{\theta}\ -Po+1)^2}$$

$$= \frac{-P^2oe^{\theta} - Poe^{\theta}}{(Poe^{\theta} - Po + 1)^2}$$

$$= \frac{Poe^{\theta}(-Po - 1)}{(Poe^{\theta} - Po + 1)^2}$$

Putting together F_1 and F_2after differentiation to give

$$\frac{\partial^2 l(p)}{\partial \theta^2} = S\left\{\frac{-(1 - Po)Poe^{\theta}}{(Po\, e^{\theta} - Po + 1)^2}\right\} + (N - S)\frac{Poe^{\theta}(-1 - Po)}{(Poe^{\theta} - Po + 1)}$$

$$for\ \frac{\partial^2\, l(p)}{\partial \theta^2}$$

at $\theta = 0$

$$= S\left(\frac{-Po(1 - Po)}{1^2}\right) + (N - s)\frac{Po(-1 - Po)}{1^2}$$

$$= S(-Po + P^2o) + (N - S)(-Po - P^2o)$$

$$= -SP_0 + SP_0^2 + N(-P_0 - P_0^2) - S(-P_0 - P_0^2)$$

$$= NP_0 - NP_0^2 = NP_0(1 - P_0) = V$$

At this point I can happily use my reduced models (formulas) to determine the log rank statistics Z, and the null variance V that are very necessary to analyze any interim results sequentially i.e.

$$Z = S - NP_0 \text{.....equation (i)}$$

$$V = NP_0(1 - P_0)\ldots\ldots\ldots\ldots\text{equation (ii)}$$

Further I need to set stopping boundary lines, upper and lower boundaries such that when the plotted points cross the boundaries the trial can stop and decision be made. The plotted points shall take the form $(Z_1V_1), (Z_2V_2), \cdots\cdots, (Z_nV_n)$ where V values will take the x –axis and Z will be on y axis.

3.3 The Stopping Boundaries

There are two approaches of working out stopping boundaries:

The boundaries of the test are computed given type 1 $error\ \alpha$ and the power $1 - \beta$ under H_a:

1. The equations of the boundaries are given by;

Upper boundary; $Z_{UB} = a + \gamma\mu$

Lower boundary: $Z_{LB} = -a + 3\ \gamma v$

Where $a = a^1 - 0.583\sqrt{l}$

Where l denotes the increament in v between two aanakysis when discrete analyses are performed every ‘n’ patients and

$$l = nPo(1-Po)$$

$$a^1 = \frac{2}{\theta_a}\log\left(\frac{1}{2\alpha}\right)$$

and

$$\gamma = \frac{1}{4}\theta a^1$$

2. Noting the significance level by $\propto$, reference improvement by θ_R and the required power as (1-$\propto$) then:

$$Z \in (-a+\gamma V, a+\mu V) \text{ where}$$

$$\gamma = \frac{3\theta_R}{4}, \quad \mu = \frac{\theta_R}{4}, \quad a = \left(\frac{2}{\theta_R}\right)\log\left(\frac{1}{2\propto}\right).$$

Illustration

If we choose Pa^1, the smallest event rate for which further investigations are worth while, we can specify.

$$H_a: P > P_a^1$$

3.4 Flexible Stopping Methodology

The Investigator can alternatively decide to use flexible sopping approach where need not to necessary plot the graph but can continuously compare the Z and V statistics as they are sequentially obtained as shown below.

Denote the values of the statistics Z and V collected at i^{th} inspection by Z_i, and V_i respectively, and letV_0. At the i^{th} inspection, the test will terminate with the new treatment is superior if:

$$Z_i = 0.583\sqrt{V_i - V_{i-1}} > a + \mu V_i$$

And it will terminate and accept the null hypothesis if

$$Z_i = 0.583\sqrt{V_i} - V < -a + \gamma V_{\text{i}}$$

Otherwise it will continue.

Incase $\alpha \neq \beta$ based on the assumption that Z is normally distributed with mean $\theta v \; and \; variance \; V \; we \; can \; use.$

θa^1 a corrected value for θa, given by the approximate formula

$$\theta^1 a = \theta a[2\emptyset^{-1}(1-\alpha)/\{\emptyset^{-1}(1-\alpha)\alpha + \emptyset^{-1}(1-\beta\}]$$

Where $\emptyset(x)$ denotes the standard normal distribution function

$$\emptyset(x) = p(X \leq x) = \int_0^x f(x)dx$$

x~ N (O, 1)

The above fully describes a triangular Test which aims to detect any improvement of any increase of the measured endpoint response rate

CHAPTER FOUR - ANALYSIS

4.1 Design of the Triangular Test and the Results of Sequential Analysis on Simulated Data.

Table 1: A Case Study

ANALYSIS NUMBER	S	N	P	Z	V
1	0	0	0	0	0
2	5	10	0.5	0	2.5
3	10	15	0.67	2.5	3.75
4	16	20	0.8	6	5
5	21	25	0.84	8.5	6.25

4.1.1 Calculations for the Z and V Statistics

Using the equations to get the relationship:

$$Z = S - NP_0 \text{ and } V = NP_0(1 - P_0)$$

Where:

Z: is the log-rank statistic. (Benefit as compared with the null hypothesis-mean difference-efficient score). And

V: is the null variance. (Accumulated fishers' information since the beginning of the trial-parameter of interest)

S: is the number of patients who experience a response (success)

N: is the number of patients included in the study.

P_0: is the threshold value i.e. (a clinical improvement) the largest response for which further investigations is not worthwhile.

4.1.2 Assumption of the Analysis

1. A single primary response of interest that is body temperature maintaining at ≤ 37 for 48 hours.
2. Recruitment of patients was done at once and the follow up was to be after a short time of two months
3. Efficacy was the only main objective of study.
4. Z will be normally distributed with mean θV and variance V.

Calculations:

Given -

$$P_0 = 0.5$$

When s = 0, N = 0

$$Z_0 = -NP_0 \quad and \; V_0 = NP_0(1 - P_0)$$

$$Z_0 = 0 - 0 \times 0.5 = 0$$

$$V_0 = 0 \times 0.5(1 - 0.5) = 0$$

When s = 5, N = 10

$$Z_1 = 5 - 10 \times 0.5 = 0$$

$$V_i = 10 \times 0.5(1 - 0.5) = 2.5$$

When S=10, N=15.

$$Z_2 = 10 - 15 \times 0.5 = 2.5$$

$$V_2 = 15 \times 0.5(1 - 0.5) = 3.75$$

When S=16, N=20

$$Z_3 = 16 - 20 \times 0.5 = 6$$

$$V_3 = 20 \times 0.5(1 - 0.5) = 5$$

When S=21, N=25

$$Z_4 = 21 - 25 \times 0.5 = 8.5$$

$$V_4 = 25 \times 0.5(1 - 0.5) = 6.25$$

When S=24, N=30

$$Z_5 = 24 - 30 \times 0.5 = 9$$

$$V_5 = 30 \times 0.5(1 - 0.5) = 7.5$$

Each analysis has given a pair of values of V and Z statistics. The corresponding points are plotted on the graph as they become available. The plotting of Z against V is referred to us as a sample path. The horizontal axis corresponds to the V and vertical axis corresponds to Z.

4.1.3 Stopping Boundary Lines.

Given by the following boundary equations:

Z=a+μV (upper boundary line)

Z= - a+γV (lower boundary line)

Where: $\gamma = \frac{3\theta_R}{4}$

$$\mu=\frac{\theta_R}{4},$$

$$a=(\frac{2}{\theta_R})\log(\frac{1}{2\propto}).$$

In this case study we let θ_R=0.75(reference improvement)

$$\alpha=0.05$$

4.1.4 Computations

$$\gamma=3*0.75/4=0.563.$$

$$\mu=0.75/4=0.188$$

$$a=(\frac{2}{0.75})\log(\frac{1}{2}\times 0.05)=6.14.$$

Hence the upper straight stopping boundary line will be:

$$Z_{UB}=a+\mu V$$

Z_{UB}=6.14+0.188V and

Lower stopping boundary:

$$Z_{LB}=-a+\gamma V$$

Z_{LB}=-6.14+0.563V

4.2 Summary by Graph

Design of the triangular test (upper boundary=6.14+0.188V, lower boundary ,Z=-6.14+0.563V) .The Z and V statistics were calculated each time apatient responded .The corresponding points were plotted on the graph and compared with stopping boundaries on the fourth analysis ,the upper boundary was crossed causing the inclusion to be stopped.

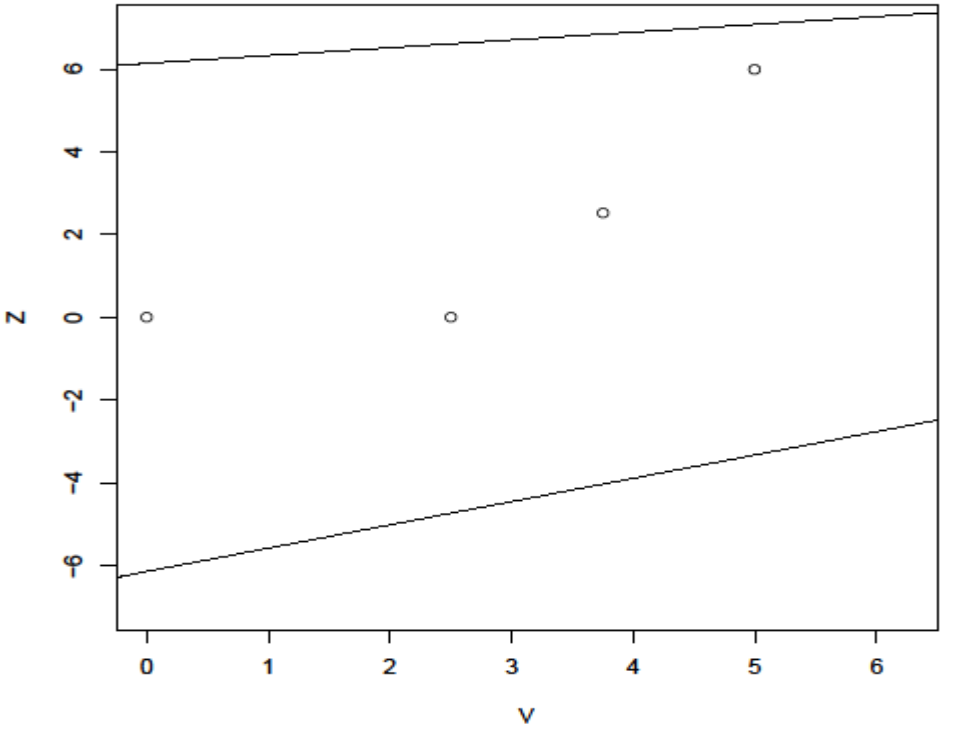

Figure 1: A Graph of V against Z and the Stopping Boundary Lines

CHAPTER FIVE - CONCLUSION AND RECOMMENDATIONS

The study was to be carried out on thirty (30) patients (fixed sample).Because the upper boundary had been crossed at the fourth inspection, it was apparent that the drug was effective at the 5% significance level (one sided alternative).Twenty one out of thirty patients (21 out of 30) had been entered into the study and so a substantial saving in sample size and time has been achieved.

Many forms of trial are not suitable for existing sequential methods because of cases in which it impossible to identify a single primary response of interest, cases in which recruitment of patient's takes place over a short time but follow up is long and in cases the comparison of efficacy is but one of several important trial objectives.

However there are many trials for which sequential analysis is suitable and desirable but is not popularly applied just because of extra work involved in monitoring the study, unfamiliarity of both medical investigators and medical statisticians with the methodology and lack of suitable software.

Basing my discussion on the above analysis and reasons given above, I appeal for additional effort s to enhance the methodology and in particular those trials that involve life- threatening diseases that need quick medical intervation.

Further I recommend to pharmaceutical companies and medical research institutions to recognize and allow sequential analysis, develop data gathering electronic systems so as to potentially benefit from the methodology.The benefits will have positive economic and ethical implications.

REFERENCES

O'Brien P.C. and Fleming T.R (1979): A multiple testing procedure for clinical trials, *Biometrics,* **35**(5) pp549-556

Pocock S. J., (1977): Clinical Trials with Multiple Outcomes: A Statistical Perspective on their Design, Analysis, and Interpretation, Elsevier Science Inc, **18** pp 530-545

Kilpatrick G.S, and Oldham P.D., :(1954) Sulphonamide prophylaxis in chronic bronchitis; a clinical trial. Br Med J. **14; 2**(4884) pp385–388

Whitehead J.P., (1983): The design and analysis of sequential clinical trials. Ellis Harwood publishers Chichester

Armitage P., McPherson and Rowe B. C. (1969): Repeated significance Tests on acuumulaating data, *STOR,* **132**(2) pp 235-244

Ian D. C., Stephen J. F. and Philip J. R., (1982): Effect of the fatty acid oxidation inhibitor 2-tetradecylglycidic acid on pyruvate dehydrogenase complex activity in starved and alloxan-diabetic rats, *Biochem,* **208** pp53–60

APPENDICES

R-code used to plot the graph.

```
Plot (x, y, ylim=c(-7,-7))      #to plot the graph with limits.

abline(a=-6.14,b=0.188)     #to plot lower boundary line.

abline(a=6.14,b=0.563)      #to plot upper boundary line.

Z=c(0,0,2.5,6,8.5)              #to plot the Z values

V=c(0,2.5,3.75,5,6.25)      #plot the v values.
```