Peter Armitage's pioneering work:
Laying the foundations for
sequential medical trials

Christopher Jennison

Department of Mathematical Sciences,

University of Bath, UK

http://people.bath.ac.uk/mascj

21st Armitage Workshop and Lecture

Cambridge

November 2024

Plan for this talk

1. From industrial quality control to sequential clinical trials
2. Armitage's testing boundaries

   Mathematical analysis   Computation

3. Anscombe's critique

   Bayesian methods   Decision theory

   Horizon problems   Response adaptive randomisation

4. Subsequent developments

   Group sequential tests,   Distribution theory,

   Error spending designs,   Optimisation,

   Inference on termination,   Pipeline data, ...

5. Promotion of sequential methods

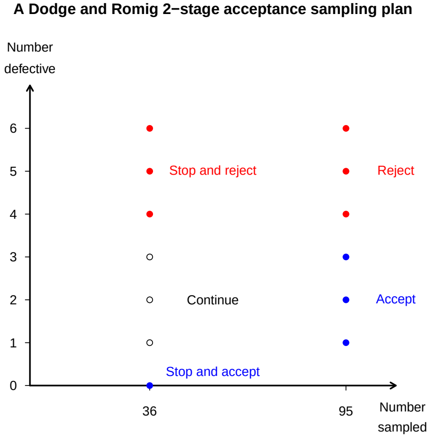
   Books and Software   Working with practitioners

Peter Armitage 1924–2024

1. Industrial Quality Control: Sequential Sampling Plans

"Sequential sampling" originated in the early 20th century.

Dodge & Romig (*Bell System Technical Journal*, 1929) presented a formulation for two-stage acceptance sampling plans.



Industrial Quality Control: Sequential Sampling Schemes

Sequential methods were developed during World War 2 for quality control in the manufacture of munitions and for comparing the effectiveness of different operational strategies.

In the USA, Abraham Wald developed the Sequential Probability Ratio Test (*JASA*, 1945).

In the UK, George Barnard developed similar methods in the SR17 unit in the Ministry of Supply, publishing some of these methods in the paper "Sequential tests in industrial statistics" (Barnard, *JRSS, Supplement*, 1946).

Peter Armitage started the Mathematical Tripos at Cambridge in 1941, but interrupted his studies to join the SR17 unit in 1943. He worked with George Barnard and would have seen sequential methods during this time. He returned to Cambridge in 1945 to complete his degree in Mathematics.

Randomised Clinical Trials

Randomised clinical trials first appeared in the UK in the 1940s.

In 1943-44, the Medical Research Council (MRC) UK conducted a trial of patulin treatment for the common cold. This trial used an alternation procedure, rather than randomisation, to allocate subjects to study groups.

Sir Austin Bradford Hill was the statistician for the 1946 trial of streptomycin as a treatment for pulmonary tuberculosis. Patients were allocate to treatment groups randomly and, after this trial, randomisation became standard practice in clinical trials.

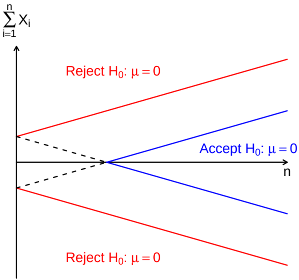
After the war, Peter Armitage worked at the Medical Research Council's Statistical Research Unit in the London School of Hygiene and Tropical Medicine. One part of his research activity was to bring sequential methods into the domain of clinical trials.

Sequential Clinical Trials

Wald's Sequential Probability Ratio Test has an elegant construction — only a simple calculation is needed to create a test with specified type I and type II error rates.

Suppose observations $X_i \sim N(\mu, \sigma^2)$, $i = 1,\, 2,\, \ldots,$ are to be observed and we wish to test H_0: $\mu = 0$ against $\mu \neq 0$.

A two−armed Sequential Probability Ratio Test



The SPRT is an "open" test with no upper limit on sample size.

2. Sequential Clinical Trials

In a medical trial one would wish to have both a small average sample size and a small maximum sample size.

Armitage (*Biometrika*, 1957) proposed "Restricted Sequential Procedures" to achieve this goal.

With a normally distributed response, some neat mathematics was needed to make calculations feasible.

Suppose observations $X_i \sim N(\mu, \sigma^2)$, $i = 1, 2, \ldots,$ are to be observed.

We wish to test H_0: $\mu = 0$ against $\mu \neq 0$ with two-sided type I error probability 2α and power $1 - \beta$ at $\mu = \pm\mu_1$.

Here, each observation could be the difference in responses for a pair of subjects on two treatment arms.

In fact, the test is applicable for unpaired data.

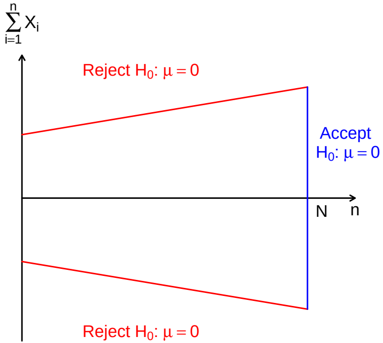
Armitage's Restricted Sequential Procedure

A Restricted Sequential Procedure has upper and lower boundaries

$$\sum_{i=1}^{n} X_i \;=\; a + b\,n \quad \text{and} \quad \sum_{i=1}^{n} X_i \;=\; -a - b\,n$$

and truncation at $n = N$.

A Restricted sequential procedure



Armitage's Restricted Sequential Procedure

Armitage was able to define a Restricted Sequential Procedure with two-sided type I error probability 2α and power $1-\beta$ at $\mu = \pm\mu_1$.

To achieve this, he applied a result for diffusion processes proved by Maurice Bartlett (1946) and a likelihood ratio argument similar to that used by Wald to define the Sequential Probability Ratio Test.

Armitage set

$$a = \frac{\sigma^2}{\mu_1} \log\left(\frac{1-\beta}{\alpha}\right), \quad b = \frac{\mu_1}{2}$$

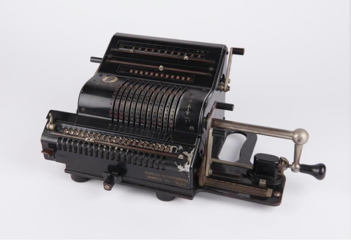
and found N satisfying

$$\beta = \Phi\left(\frac{a}{\sigma\sqrt{N}} - \frac{b\sqrt{N}}{\sigma}\right) - \left(\frac{1-\beta}{\alpha}\right)\Phi\left(\frac{-a}{\sigma\sqrt{N}} - \frac{b\sqrt{N}}{\sigma}\right).$$

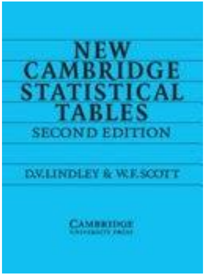
Note: Modest calculation is required, using tables of the standard normal CDF Φ.

Calculating machines

The Braunsviga Calculating Machine



Cambridge research students used to spend time each week doing the calculations needed to produce statistical tables.



NEW CAMBRIDGE STATISTICAL TABLES

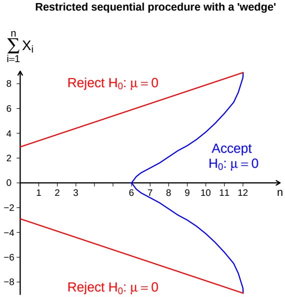
SECOND EDITION

D.V. LINDLEY & W.F. SCOTT

CAMBRIDGE UNIVERSITY PRESS

Armitage's Restricted Sequential Procedure

Schneiderman & Armitage (*Biometrika*, 1962) inserted a "wedge" in the continuation region to facilitate early stopping for $\mu \approx 0$.



The boundary for the "wedge" was derived analytically but this still required significant computation.

Average sample sizes were computed by Monte Carlo simulation with 100 replicates

Calculations were carried out using the NIH's IBM 650 computer.

IBM 650 computer

The IBM 650

The IBM 650's
Magnetic memory drum

"The average time for accessing data or programming was 2.4 milliseconds, less than the time it takes a fruit fly to flap its wings"

Compare current terminology: A petaflop is one quadrillion (10^{15}) floating-point operations per second.

Repeated Significance Tests

There had been debate on whether it was appropriate to analyse accumulating data repeatedly without making an "adjustment" for the repeated analyses.

Armitage, McPherson & Rowe (*JRSS, A*, 1969) made precise calculations of the probability of a type I error if one conducts a sequence of K two-sided significance tests, each at level 2α.

Suppose $X_i \sim N(\mu, 1)$, $i = 1,\, 2,\, \ldots,$ are observed,

$$S_n = \sum_{i=1}^{n} X_i, \quad n = 1, \ldots, K,$$

and the null hypothesis H_0: $\mu = 0$ is rejected after observation n if

$$|S_n| \geq \sqrt{n}\, \Phi^{-1}(1 - \alpha).$$

How does the overall probability of a type I error grow with n?

Repeated Significance Tests

Consider the sequential test which stops to reject H_0 after observation n if

$$|S_n| \geq c_n = \sqrt{n}\,\Phi^{-1}(1-\alpha).$$

Let $f_n(s_n)$ be the probability density function of S_n.

Armitage, McPherson & Rowe noted that, under H_0: $\mu = 0$, f_1 is the standard normal density and f_n is related to f_{n-1} by the recursive formula

$$f_n(s_n) = \begin{cases} \int_{-c_{n-1}}^{c_{n-1}} f_{n-1}(u) \frac{1}{\sqrt{2\pi}} \exp\{\frac{-(s_n-u)^2}{2}\}\, du, & -c_n \leq s_n \leq c_n \\ 0 & \text{otherwise} \end{cases}$$

They applied numerical integration to calculate the probabilities $P(|S_n| > c_n)$ for $n = 1, \ldots, K$.

Computations were made on the Institute of Computer Science's Atlas computer and University College London's IBM 360/65.

Repeated Significance Tests

Suppose a test of H_0 is carried out at two-sided significance level $\alpha = 0.05$ on K occasions during the course of the trial.

Armitage, McPherson & Rowe found the *overall* type I error rate:

Number of tests, K	Overall error rate	Number of tests, K	Overall error rate
1	0.050	10	0.193
2	0.083	20	0.248
3	0.107	100	0.374
4	0.126	200	0.424
5	0.142	∞	1.000

They also reported results from Monte Carlo simulations but observed that the recursive formulae for the densities $f_n(s_n)$ combined with numerical integration gave more accurate answers.

Repeated Significance Tests

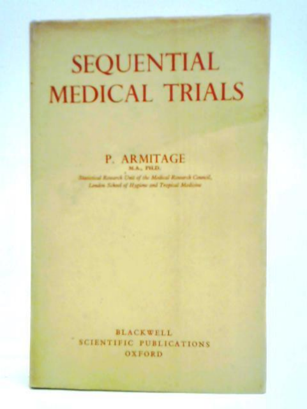
Armitage, McPherson & Rowe noted that, when a fixed number of tests is specified, these can be conducted at a level α' chosen so that the overall type I error probability is a desired value α.

Values of α' to achieve $\alpha = 0.05$ are:

Number of tests, K	Critical value for each Z_n	Nominal two-sided significance level, α'
1	1.960	0.0500
5	2.413	0.0158
10	2.555	0.0106
20	2.672	0.0075
50	2.797	0.0052
100	2.875	0.0040
200	2.941	0.0033

Repeated Significance Tests

Armitage published the book “Sequential Medical Trials” in 1960.



SEQUENTIAL
MEDICAL TRIALS

P. ARMITAGE
M.A., PH.D.
Statistical Research Unit of the Medical Research Council,
London School of Hygiene and Tropical Medicine

BLACKWELL
SCIENTIFIC PUBLICATIONS
OXFORD

This was followed by a second edition in 1975 which placed a greater emphasis on repeated significance tests.

Repeated Significance Tests

The methods developed by Armitage were applied.

In the *Annual Review of Medicine* (1969), Armitage wrote:

> *Sequential trials have been reported in over 50 papers . . . . The branches of medicine and surgery involved are varied . . . cardiovascular or cerebrovascular disease . . . respiratory disease . . . the nervous system . . . tetanus.*

Gehan & Schneiderman (*Statistics in Medicine*, 1990) summarised the outcome of a trial comparing treatments for childhood leukemia:

> *The minimum number of pairs of patients was 9 and the maximum number was 66. The trial reached a sequential boundary favouring 6MP after 18 preferences had occurred, 15 for 6MP and 3 for placebo.*

3. Anscombe on Armitage's "Sequential Medical Trials"

In 1963, Frank Anscombe wrote an article in *JASA* which he described as "statistical polemic thinly disguised as a book review". In the summary, Anscombe states

> *"Sequential analysis is a hoax"*

Nevertheless, he starts by saying

> *"Before any adverse criticisms . . . it is proper to make two observations in defense of the book. First, its net effect on medical research will almost certainly be good . . . "*

In his reply (*JASA*, 1963) Armitage wrote

> *"The present note could be regarded as a reply thinly disguised as statistical polemic, for I certainly do not wish to strike an aggressive attitude."*

Armitage and Anscombe had been friends since working together at the SR17 unit during the war.

Anscombe on Armitage's book "Sequential Medical Trials"

Anscombe criticised Armitage's focus on frequentist error rates and advocated a Bayesian approach to optimise the treatment of patients within the trial and external to it.

He formulated a "horizon" problem in which it is assumed N patients with a disease are to receive one of two treatments; see also Colton (*JASA*, 1963).

After the trial involving $2n$ of the N individuals, the remaining $N - 2n$ patients will receive the treatment selected as the better at the conclusion of the trial.

Anscombe compared the sequential stopping boundaries for such a trial with those proposed by Armitage.

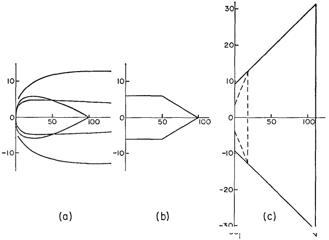


FIGURE 1. Stopping boundaries for the comparison of two treatments. The abscissa is n, the number of pairs of patients. The ordinate is y, the cumulative sum of response differences. (a) Three boundaries from Table 1. (b) A simplified boundary. (c) Two boundaries given by Armitage.

Anscombe on Armitage's book "Sequential Medical Trials"

In his reply, Armitage challenged the practicality of Anscombe's methods. He also questioned the ethics of optimising treatment of future patients at the expense of current patients.

I would note that now, in order to satisfy regulators in a New Drug Application, the role of a typical Phase III trial is one of testing rather than treatment selection.

The "horizon problem" formulation became an attractive setting for Response Adaptive Randomisation rules — but issues of practicality remained. See, for example, John Anderson's comments on Bather's read paper (*JRSS, B*, 1981)

In a clinical trial for a rare disease, the participants in a clinical trial may comprise a large part of the patient population.

In such cases, regulators may be willing to consider treatment selection, as opposed to testing with a stringent requirement on the P-value.

4. What next: Group Sequential Tests

In June 1965, a group of statisticians experienced in clinical trials gathered at an NIH seminar to discuss the role of hypothesis testing in clinical trials; see Cutler at al. (*Journal of Chronic Diseases*, 1965).

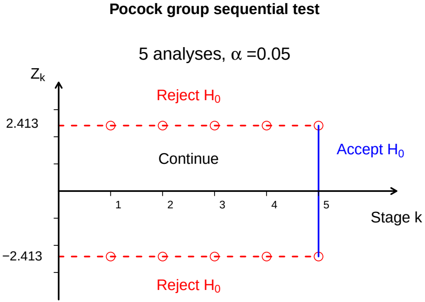
In the discussion, Lawrence Shaw (Veterans Administration) proposed conducting a small number of interim analyses, using the term "block sequential analysis".

Over a decade later the papers by Pocock (*Biometrika*, 1977) and O'Brien & Fleming (*Biometrics*, 1979) prompted the widespread adoption of group sequential tests.

Group sequential methods were well-suited to a multi-centre trial with an Independent Data Monitoring Committee and a separate party (e.g., a Clinical Research Organisation) cleaning and analysing the accruing data.

Group Sequential Tests

Stuart Pocock (a research student of Armitage) suggested applying a repeated significance test with a small number of analyses



He used the numerical methods of Armitage, McPherson & Rowe.

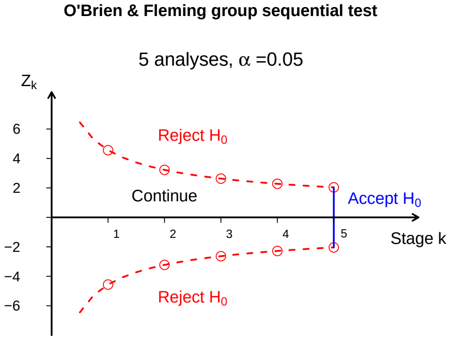
He showed how to choose group sizes to achieve a specified power.

He demonstrated protection of the overall type I error rate for a variety of response distributions and test statistics.

Group Sequential Tests

O'Brien & Fleming suggested a different group sequential test.

Their boundary is wide early on and the final critical value for the Z-statistic is only slightly higher than of a fixed sample size test.



They used simulation with 10,000 replicates to define their boundaries and reported properties based on 1,000 simulations.

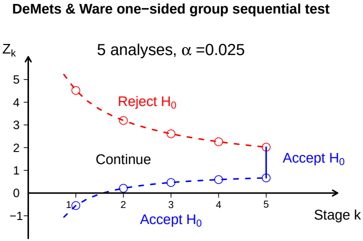
One-sided Group Sequential Tests

DeMets & Ware (*Biometrika*, 1980, 1982) noted that in a comparison of a new treatment to a control, the hypothesis testing formulation should be one-sided.

If θ is the improvement from using the new treatment (the "treatment effect"), we should test H_0: $\theta \leq 0$ against $\theta > 0$.

If $\theta \leq 0$, it would be unethical to randomise patients in order to learn whether $\theta = 0$ or $\theta < 0$.

DeMets & Ware proposed one-sided group sequential tests.



Joint Distribution Theory

The primary endpoint in a clinical rial may be binary, normal, time-to-event, or follow some other parametric model.

In many cases, large sample theory tells us that, approximately,

$$\widehat{\theta} \;\sim\; N(\theta, \mathcal{I}^{-1})$$

where $\mathcal{I}$ is the "Fisher information" for θ.

In a trial with K analyses, let $\widehat{\theta}_k$ be the estimate of θ at analysis k.

In many cases, $\widehat{\theta}_1, \ldots, \widehat{\theta}_K$ are approximately multivariate normal,

$$\widehat{\theta}_k \sim N(\theta, \mathcal{I}_k^{-1}), \quad k = 1, \ldots, K,$$

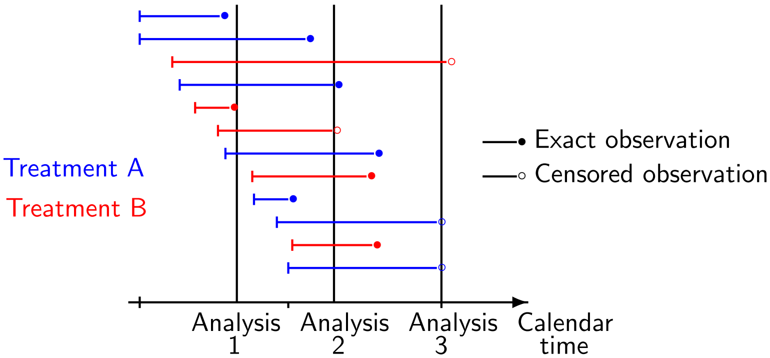
and

$$\mathsf{Cov}(\widehat{\theta}_{k_1}, \widehat{\theta}_{k_2}) = \mathsf{Var}(\widehat{\theta}_{k_2}) = \mathcal{I}_{k_2}^{-1} \quad \text{for } k_1 < k_2.$$

This "canonical joint distribution" was proved in considerable generality by Jennison & Turnbull (*JASA*, 1997) and Scharfstein, Tsiatis & Robins (*JASA*, 1997).

Survival Data

Time-to-event data pose particular problems for sequential analysis.



Subjects are randomised to a treatment as they enter the study.

Survival is measured from entry to the study.

At an interim analysis, subjects are censored if they are still alive.

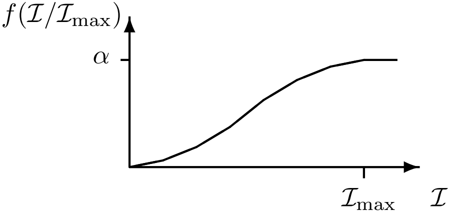
Information at a specific calendar time is unpredictable.

Error Spending Tests

When the sequence $\mathcal{I}_1$, $\mathcal{I}_2$, ... is unpredictable, a group sequential design must adapt to observed information levels.

Lan & DeMets (*Biometrika*, 1983) introduced "error spending" tests of H_0: $\theta = 0$ against $\theta \neq 0$.

Maximum information design with spending function $f(\mathcal{I}/\mathcal{I}_{\max})$



The boundary at analysis k is set to give cumulative type I error probability $f(\mathcal{I}_k/\mathcal{I}_{\max})$.

If $\mathcal{I}_{\max}$ is reached without rejecting H_0, then H_0 is accepted.

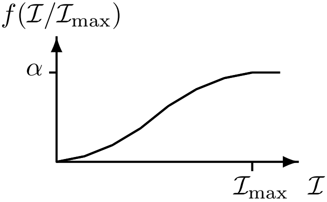
Error Spending Tests

For a one-sided test of H_0: $\theta \le 0$ against $\theta > 0$ with

Type I error probability α at $\theta = 0$,

Type II error probability β at $\theta = \delta$,

we need two error spending functions.



Type I error probability α is spent according to the function $f(\mathcal{I}/\mathcal{I}_{\max})$, and type II error probability β according to $g(\mathcal{I}/\mathcal{I}_{\max})$.

Error Spending Tests

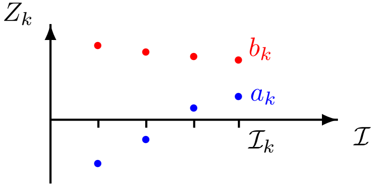
Analysis k: Observed information $\mathcal{I}_k$

Find a_k and b_k to satisfy

$$P_{\theta=0}\{a_1 < Z_1 < b_1, \ldots, a_{k-1} < Z_{k-1} < b_{k-1}, Z_k > b_k\}$$
$$= f(\mathcal{I}_k/\mathcal{I}_{\max}) - f(\mathcal{I}_{k-1}/\mathcal{I}_{\max}),$$

and

$$P_{\theta=\delta}\{a_1 < Z_1 < b_1, \ldots, a_{k-1} < Z_{k-1} < b_{k-1}, Z_k < a_k\}$$
$$= g(\mathcal{I}_k/\mathcal{I}_{\max}) - g(\mathcal{I}_{k-1}/\mathcal{I}_{\max}).$$



Optimal Stopping Boundaries

One can ask

What is the best choice of stopping boundary?

What is the best choice of error spending function?

What is needed for these questions to be well-posed?

Consider a test of H_0: $\theta \leq 0$ against $\theta > 0$ with type I error probability α and power $1-\beta$ at $\theta = \delta$.

A fixed sample size study needs information

$$\mathcal{I}_{fix} = \frac{\{\Phi^{-1}(1-\alpha) + \Phi^{-1}(1-\beta)\}^2}{\delta^2},$$

where Φ is the standard normal CDF.

We shall describe methods and results in terms of information, noting that this is (roughly) proportional to sample size in many clinical trial settings.

Optimal Stopping Boundaries

A group sequential test (GST) with K analyses will require a maximum information possible level $\mathcal{I}_K$, greater than $\mathcal{I}_{fix}$.

We call $R = \mathcal{I}_K / \mathcal{I}_{fix}$ the *inflation factor* of a group sequential test.

We can seek a GST that minimises expected information $\mathbb{E}_\theta(\mathcal{I})$ under certain values of the treatment effect, θ, with a given number of analyses K and inflation factor R.

We may aim to minimise

$$\sum_i w_i \, \mathbb{E}_{\theta_i}(\mathcal{I})$$

for selected treatment effects θ_i and weights w_i.

Alternatively, we may minimise

$$\int f(\theta) \, \mathbb{E}_\theta(\mathcal{I}) \, d\theta,$$

where f is, say, a normal density.

Computing Optimal Group Sequential Tests

In optimising a GST, Eales & Jennison (*Biometrika*, 1992) and Barber & Jennison (*Biometrika*, 2002) create a Bayes sequential decision problem, placing a prior on θ and defining costs for sampling and for making incorrect decisions.

Such a problem can be solved rapidly using the numerical integration methods of Armitage, McPherson & Rowe, combined with dynamic programming.

One then searches for the combination of prior and costs such that the solution to the (unconstrained) Bayes decision problem has the specified frequentist error rates α at $\theta = 0$ and β at $\theta = \delta$.

The resulting design solves both the Bayes decision problem and the original frequentist problem.

Although the Bayes decision problem is formed as a computational device, this derivation demonstrates that an efficient frequentist design should be a good Bayesian procedure, and vice versa.

Benefits of Group Sequential Testing

One-sided GSTs with binding futility boundaries, minimising $\{\mathbb{E}_0(\mathcal{I}) + \mathbb{E}_\delta(\mathcal{I})\}/2$ for K equally sized groups, $\alpha = 0.025$, $1 - \beta = 0.9$ and $\mathcal{I}_{max} = R\mathcal{I}_{fix}$.

Minimum values of $\{\mathbb{E}_0(\mathcal{I}) + \mathbb{E}_\delta(\mathcal{I})\}/2$, as a percentage of $\mathcal{I}_{fix}$

	R					*Minimum*
K	1.01	1.05	1.1	1.2	1.3	*over* R
2	80.8	74.7	**73.2**	73.7	75.8	73.0 at R=1.13
3	76.2	69.3	66.6	**65.1**	65.2	65.0 at R=1.23
5	72.2	65.2	62.2	59.8	**59.0**	58.8 at R=1.38
10	69.2	62.2	59.0	56.3	**55.1**	54.2 at R=1.6
20	67.8	60.6	57.5	54.6	**53.3**	51.7 at R=1.8

Note: $\mathbb{E}(\mathcal{I}) \searrow$ as $K \nearrow$ but with diminishing returns,

$\mathbb{E}(\mathcal{I}) \searrow$ as $R \nearrow$ up to a point.

Efficient Error Spending GSTs

In their book, *Group Sequential Methods with Applications to Clinical Trials*, Jennison & Turnbull (1999) suggest using the "ρ-family" of one-sided error spending tests.

A target information level $\mathcal{I}_{max}$ is specified and the type I and type II error probabilities "spent" up to analysis k are, respectively,

$$f(\mathcal{I}) = \min\{(\mathcal{I}/\mathcal{I}_{\max})^2, 1\}\,\alpha \quad \text{and} \quad g(\mathcal{I}) = \min\{(\mathcal{I}/\mathcal{I}_{\max})^2, 1\}\,\beta.$$

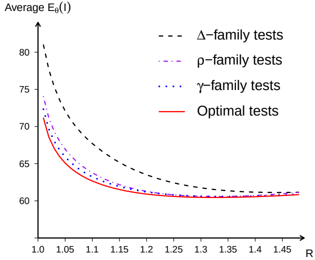
The value of ρ governs the rate at which error probability is spent, with $\rho = 1$ producing Pocock-type boundaries and $\rho = 3$ producing O'Brien & Fleming-type boundaries.

The choice of ρ determines the inflation factor R and $\mathcal{I}_{max}$ is R times the information needed for a fixed sample test.

Barber & Jennison (2003) show this family yields tests that are close to optimal for a variety of measures of $\mathbb{E}_\theta(\mathcal{I})$.

Efficient Error Spending GSTs

Plots of $\int f(\theta) E_\theta(\mathcal{I}_T) d\theta$ as a percentage of fixed sample $\mathcal{I}_{fix}$ vs inflation factor R for tests with 5 analyses, $\alpha = 0.025$ and $\beta = 0.1$.



Here, $f(\theta)$ is the density of a $N(\delta, \delta^2/4)$ distribution and $\mathcal{I}_{fix}$ the information required for a fixed sample size test.

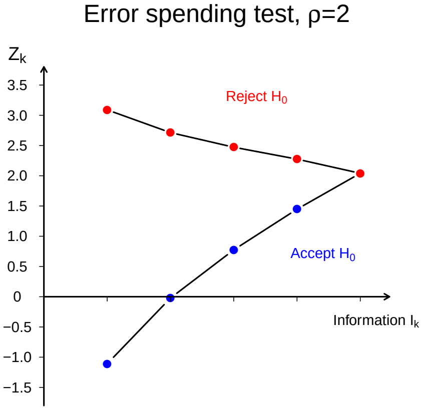
The Δ family of parametric boundaries is that proposed by Wang & Tsiatis (*Biometrics*, 1987).

The Gamma family of error spending functions is as described by Hwang, Shih & De Cani (*Statistics in Medicine*, 1990).

Efficient Error Spending GSTs

A test with 5 planned analyses, type I error probability $\alpha = 0.025$, power 0.9 if $\theta = \delta = 1$, and type I and II error spending functions

$$f(\mathcal{I}) \;=\; \min\{(\mathcal{I}/\mathcal{I}_{\max})^2, 1\}\,\alpha, \;\; g(\mathcal{I}) \;=\; \min\{(\mathcal{I}/\mathcal{I}_{\max})^2, 1\}\,\beta.$$



$$R = 1.132.$$

Efficient Error Spending GSTs

Similar methods can be used to

Optimise timing of analyses,

Allow data dependent group sizes (Jennison & Turnbull, *Biometrika*, 2006).

However, there is little to be gained from these embellishments.

Other applications of this method of optimisation include:

Group sequential tests of superiority and non-inferiority (Öhrn & Jennison, *Statistics in Medicine*, 2010),

Group sequential tests that can deal with "pipeline data" (Hampson & Jennison, *JRSS, B*, 2013),

Optimising gain functions from financial models (Robbie Peck, *University of Bath, PhD thesis*, 2020).

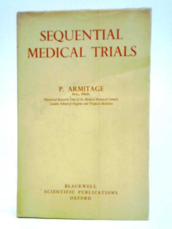
More topics

If time allowed, we could delve further into:

- Inference after a GST (see Armitage, *Biometrika*, 1958)
- Pipeline data
- Testing multiple endpoints
- Comparing multiple treatments
- Flexible designs (including Repeated Confidence Intervals)
- Adaptive designs
  - Sample size re-estimation
  - Seamless Phase 2/3 trials
  - Enrichment designs
  - Multi-arm multi-stage studies (Basket, Umbrella, . . .)

Promotion of sequential methods

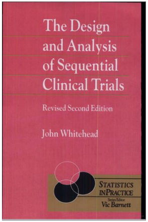
Books:



SEQUENTIAL MEDICAL TRIALS

P. ARMITAGE
M.A., PH.D.
Statistical Research Unit of the Medical Research Council,
London School of Hygiene and Tropical Medicine

BLACKWELL
SCIENTIFIC PUBLICATIONS
OXFORD

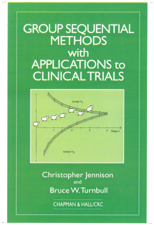


The Design and Analysis of Sequential Clinical Trials

Revised Second Edition

John Whitehead

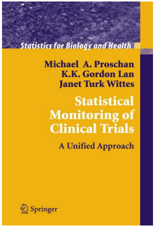
STATISTICS IN PRACTICE
Series Editor
Vic Barnett



GROUP SEQUENTIAL METHODS with APPLICATIONS to CLINICAL TRIALS

Christopher Jennison
and
Bruce W. Turnbull

CHAPMAN & HALL/CRC



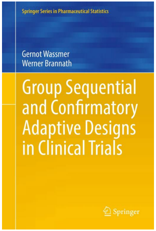
Statistics for Biology and Health

Michael A. Proschan
K.K. Gordon Lan
Janet Turk Wittes

Statistical Monitoring of Clinical Trials

A Unified Approach

Springer

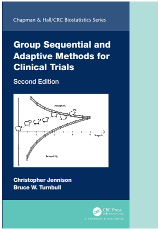


Springer Series in Pharmaceutical Statistics

Gernot Wassmer
Werner Brannath

Group Sequential and Confirmatory Adaptive Designs in Clinical Trials

Springer



Chapman & Hall/CRC Biostatistics Series

Group Sequential and Adaptive Methods for Clinical Trials

Second Edition

Christopher Jennison
Bruce W. Turnbull

CRC Press

Promotion of sequential methods

Software

PEST (Whitehead)

East (Cytel)

gsDesign (Andersen)

Addplan (Wassmer/ICON/Berry Consultants)

RPACT (Wassmer & Pahlke)

Engagement with practitioners

Collaboration

Consultancy

Societies

Statisticians in the Pharmaceutical Industry (PSI)

Society for Clinical Trials

International Society for Clinical Biostatistics, . . .

Peter Armitage 1924–2024

Peter built the foundations for sequential analysis of medical trials.

He was a statistician, a scholar and teacher.

His kind approach is a lesson to us all.

Thank you, Peter.