Institution: University of Bath

Unit of Assessment: 10: Mathematical Sciences

Title of case study: Improving Clinical Trials by Innovative Statistical Design

1. Summary of the impact

Clinical trials form a crucial step in translating fundamental medical research into improved healthcare. Many hundreds of trials are conducted every year, each involving hundreds, sometimes thousands, of patients. These trials are expensive, with costs as high as 20 or 30 thousand pounds per patient. Research in Bath on **group sequential monitoring** and the **adaptive design** of clinical trials has improved the conduct of clinical trials, leading to:

- faster results: making effective new treatments available sooner; also, stopping negative trials early avoids exposing patients to ineffective treatments and releases resources for new studies;
- smaller sample sizes: average reductions of 20-30% are possible in sequential trials;
- the ability to modify trial conditions while retaining statistical validity: this flexibility can accelerate the drug development process by months or even years.

The impact of this research is economic (the business performance of pharmaceutical companies and businesses that support them), societal (by enhancing public health and by changing the policies adopted by regulators) and ethical (ensuring clinical trials remain safe, while bringing life-saving treatments into clinical use as rapidly as possible).

2. Underpinning research

Randomised clinical trials originated in the 1940s. **Group sequential** methods, which involve monitoring results on a small number of occasions during the course of a trial, were proposed in the 1970s. The group sequential approach has steadily gained favour as:

- designs for trials have been developed to meet practical needs more closely;
- dedicated software has become available;
- clear expositions of the statistical methods and their validity have been provided.

While research groups worldwide have contributed to this field, the series of contributions by Jennison in Bath has been distinctive, influential and widely applied. The following description of the underpinning research is organised into three parts for clarity.

(i) ***Core theory and methodology for group sequential tests***

The research article [1], cited over 40 times, derives theory for the joint probability distributions that arise when accumulating data are analysed repeatedly. This allows a common treatment of sequential hypothesis tests for different types of clinical response data and provides the foundation for a unified approach to a wide variety of trial settings. Results in [1] widen the applicability of methods for simple, normally distributed responses to general parametric models incorporating baseline covariates, to survival endpoints, and to longitudinal data. This unified approach is further developed and expanded in the book [2], which applies the theory in [1] to a series of important examples. Section 7.3 of [2] extends error-spending designs to one-sided, group sequential hypothesis tests, which are now a standard requirement in many Phase III clinical trials. The stopping rule for an error spending design is produced by allocating a portion of the overall type I error probability to each analysis, depending on the observed statistical information at that point, so dealing flexibly and efficiently with the common problem of unpredictable increments in sample size and information between analyses.

(ii) ***Optimal group sequential designs and adaptive sample size modification***

Sequential monitoring aims to reduce the number of patients needed in a study. Results in [3] quantify the maximum savings that group sequential testing can achieve and knowledge of these optimal values allows the efficiency of other designs to be assessed. In particular, the Rho family of error spending tests proposed in [2] is seen to offer a highly efficient set of designs.

The last decade has seen intense interest in adaptive trial designs in which aspects of a study are altered as data are observed. An important question is whether it is actually advantageous to modify a trial's sample size adaptively in response to interim estimates of the treatment effect. Case studies and the general theory of optimal design reported in [3], [4] and [5] give a clear answer to this question: traditional, non-adaptive group sequential tests capture almost all of the efficiency gains that can be achieved by more complex adaptive designs; moreover, a poor choice of adaptive design can be substantially inferior to a good group sequential design.

(iii) ***Combination tests for survival response data in adaptive designs***

Adaptive designs enable innovation in clinical trials, such as the data-driven selection of one of several dose levels during the course of a study, or deciding whether to focus attention on a patient subgroup in which the new treatment shows a more substantial effect. The combination test of Bauer and Köhne (*Biometrics*, 1994) controls the type I error rate in a great many situations and is a key element of adaptive designs. However, it is well known that application of the combination test to survival endpoints can inflate the type I error rate. The new form of combination test for survival data defined in [6] solves this important and longstanding problem – moreover, this work has been taken up immediately by practitioners.

The above research was carried out by Jennison at Bath, where he has been Professor of Statistics since 1993. Items [1], [2], [4] and [5] were written in collaboration with Professor Bruce Turnbull (Cornell University); [3] is joint work with Stuart Barber, a Bath PhD student from 1996–99; [6] is joint work with Martin Jenkins and Andrew Stone of AstraZeneca, Macclesfield.

3. References to the research

References that best indicate the quality of the underpinning research are starred.

[1]* Jennison, C. and Turnbull, B. W. (1997) Group sequential analysis incorporating covariate information. *Journal of the American Statistical Association*, 92, 1330–1341. [doi: 10.1080/01621459.1997.10473654]

[2] Jennison, C. and Turnbull, B.W. (2000) *Group Sequential Tests with Applications to Clinical Trials*, Chapman and Hall/CRC, 390 pages. [ISBN 0-8493-0316-8]

This book has been translated into Japanese by Professors M. Toshihiko and T. Yamanaka. [ISBN 978-4-9902097-4-2]

[3] Barber, S. and Jennison, C. (2002) Optimal asymmetric one-sided group sequential tests. *Biometrika*, 89, 49–60. [doi: 10.1093/biomet/89.1.49]

[4]* Jennison, C. and Turnbull, B. W. (2003) Mid-course sample size modification in clinical trials based on the observed treatment effect. *Statistics in Medicine*, 22, 971–993. [doi: 10.1002/sim.1457]

[5]* Jennison, C. and Turnbull, B. W. (2006). Adaptive and non-adaptive group sequential tests. *Biometrika*, 93, 1–21. [doi: 10.1093/biomet/93.1.1]

[6] Jenkins, M., Stone, A. and Jennison, C. (2011) An adaptive seamless phase II/III design for oncology trials with subpopulation selection using correlated survival endpoints. *Pharmaceutical Statistics*, 10, 347–356. [doi: 10.1002/pst.472]

4. Details of the impact

(i) ***Impact of core theory and methodology on clinical trial practice***

As outlined in Section 2, the theory in [1] and its application in [2] are fundamental to a unified treatment of stopping rules for the termination of a clinical trial. Early stopping in favour of a new therapy **benefits patients** beyond the trial by making an effective new treatment available sooner and **increases the financial return to the manufacturer**. Halting a trial early when a new treatment performs poorly **releases resources** for studies of other promising therapies.

The unified theory in [1] and [2] has found wide applicability and met practical needs. The book [2] also synthesises a large volume of research into a form accessible to statisticians responsible for the design of clinical trials in the pharmaceutical industry. It has sold 3,669 copies, had over 2,100 chapter downloads (A), been translated into Japanese [2], and become the standard reference work for the industry. Google Scholar shows that since 2008, over 40 medical journal publications have cited [2] as the source of methodology used in specific clinical trials: the outcomes of these trials have affected treatment of patients in the REF period (with more rapid effect when early stopping occurred); in later examples, the trials were conducted wholly in the REF period.

We now give some specific examples of clinical trials. Publications (B) and (C) report trials that relied on [1] to design a group sequential trial with special types of response. In (B), analysis of covariance is used to adjust for baseline variables while (C) has a survival response. The report in publication (D) is of a trial using the methodology in Chapter 7 of [2] and the article states explicitly that *"Jennison and Turnbull's Rho stopping boundary (ρ = 3.0) was used".*

Economic benefit accrues also to producers of statistical software. Cytel's *East* software for the design and analysis of sequential trials draws on Jennison's work in its implementation of group sequential designs (E). In his letter (F), the president of Cytel states

"We have annual revenues of about $27,000,000. Our flagship software package East© is used by almost all major pharmaceutical companies (e.g., GSK, Novartis, Pfizer, Merck, Amgen, Lilly, Genentech), numerous smaller pharma and biotech companies … and governmental agencies (e.g., FDA, NIH). A heavily used module in East is the "Survival Module" for the design and interim monitoring of trials with mortality endpoints … The statistical methodology that we have implemented in East for such trials relies on the theory that was published by Jennison and Turnbull (JASA, 1997). This seminal paper has had a huge impact on clinical trials and has facilitated the use of group sequential and adaptive methods that can save patient resources and bring new drugs to market faster. It is fair to say that many companies have purchased East almost entirely because of its Survival Module. The reason that the methodology developed by Jennison and Turnbull (JASA, 1997) has been so influential is that it provides a unified group sequential theory that covers normal, binomial and survival distributions, with or without covariates."

(ii) ***Impact of research into optimal group sequential and adaptive design on FDA policy***

Research at Bath has shaped the policy of the US Food and Drug Administration (FDA). Jennison's results on the optimisation of clinical trial designs for particular objectives have informed regulators who make policy for pharmaceutical companies to follow. The "Guidance for Industry" of February 2010 (G) describes **FDA policy** on adaptive designs in Phase III trials. This document cites [2], [4] and two further articles by Jennison which draw on results in [3] and [5]. The conclusions in [3] and [5] are re-iterated in the statement (G, Section VI C):

" … one adaptive design approach is to allow an increase in the initially planned study sample size based on knowledge of the unblinded treatment-effect sizes at an interim stage … In general, using this approach late in the study is not advisable … The potential to decrease the sample size is best achieved through group sequential designs with well-understood alpha spending rules … ."

The Associate Director for Adaptive Design and Pharmacogenomics in the FDA's Office of Biostatistics, writes (H)

> *"Jennison's work has been invaluable in providing benchmarks by which to judge group sequential designs, in appraising the benefits of novel proposals for adaptive designs, and in extending adaptive methods to overcome impediments to their application."*

(iii) ***Use of a new combination test for survival data in an adaptive clinical trial***

The new form of combination test defined in [6] provides a sound basis for adaptive designs with survival endpoints and guarantees the crucial property, insisted upon by regulators, that the type I error rate is controlled unequivocally. Previous proposals failed to do this. The impact of this research output has already been seen in Hoffman-La Roche's GATSBY trial (study BO27952), a multi-national trial of treatments for advanced gastric cancer. The adaptive design of this Phase III trial relies fundamentally on results in [6]. The trial starts by comparing two treatment formulations against a control then, at an interim analysis, just one of these formulations is chosen for comparison with the control in the remainder of the study. Combining treatment selection and testing in a single trial achieves the statistical requirements **with fewer patients**. Our new form of combination test is used to analyse data from the two phases of the trial in a way which guarantees full control of the type I error rate. The letter from Hoffman-La Roche's Associate Director of Biostatistics (J) quotes the reference to [6] in the trial protocol:

> *"As discussed in Jenkins et al. (2011), as the extent of follow up of Stage 1 patients remains unchanged, the final testing procedure described within the Roche study BO27952 guarantees full control of the Type I error rate."*

Communication of research findings

Jennison has enhanced the impact of his research by communicating results to practitioners. Since 2008, he has presented 10 short courses on group sequential and adaptive methods, from half a day to 2 days in length, at conferences and to companies. He speaks frequently at conferences with a high proportion of industrial participants; see the listing of around 10 talks per year at http://people.bath.ac.uk/mascj. He also provides consultancy to companies on the design of individual trials with innovative group sequential or adaptive features.

5. Sources to corroborate the impact

(A) Letter from Senior Acquisitions Editor for Statistics, Chapman & Hall/CRC, Taylor & Francis.

(B) Lo, A. C. et al, (2010) Robot-assisted therapy for long-term upper-limb impairment after stroke, *New England Journal of Medicine*, 362, 1772–1783. [doi: 10.1056/NEJMoa0911341]

(C) Loehrer, P.J. et al (2011) Gemcitabine alone versus gemcitabine plus radiotherapy in patients with locally advanced pancreatic cancer: An Eastern Cooperative Oncology Group trial, *Journal of Clinical Oncology*, 31, 4105–4112. [doi: 10.1200/JCO.2011.34.8904]

(D) Barrios, C.H. et al (2010) Phase III randomized trial of sunitinib versus capecitabine in patients with previously treated HER2-negative advanced breast cancer, *Breast Cancer Research and Treatment*, 121, 121–131. [doi: 10.1007/s10549-010-0788-0]

(E) East5 manual, p.1255: refers to "Rho spending function … Jennison and Turnbull (2000)" [available from the HEI].

(F) Letter from President of Cytel Inc. (producers of the *East* software package).

(G) U.S. FDA, "Guidance for Industry Adaptive Design Clinical Trials for Drugs and Biologics", February 2010 http://www.fda.gov/downloads/Drugs/.../Guidances/ucm201790.pdf [also available from the HEI].

(H) Letter from Associate Director for Adaptive Design and Pharmacogenomics, U.S. FDA.

(J) Letter from Associate Director Biostatistics, F. Hoffman-La Roche Ltd.