Adaptive Designs:

New options for clinical trials

Christopher Jennison

Department of Mathematical Sciences,

University of Bath, UK

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

16th CR@B Symposium

November 2017

Role of clinical trials in oncology

For cytotoxic drugs, trials with human subjects:

- Phase I: Determining the Maximum Tolerated Dose
- Phase II: To confirm potential efficacy
- Phase III: Confirmatory trials

Phase III trials: Sample size calculations

The randomised clinical trial compares a new treatment against a control treatment, typically the current "standard of care".

A type I error probability (false positive rate) is specified.

The sample size is chosen to guarantee a high probability (power) of showing the new treatment to be superior when the difference between this treatment and control is a certain size.

When the primary endpoint is overall survival, meeting the power condition will require a certain number of observed deaths.

Phase III trials: Early stopping

Stopping for a positive outcome

If the treatment difference is large, the superiority of the new treatment may become evident early in the trial.

Then, an early decision in favour of the new treatment is desirable:

- To reduce the cost of the trial,
- To avoid randomising more patients to an inferior treatment,
- To hasten introduction of the new treatment into general use.

Stopping for a negative outcome

Analysis of interim data may show that it is unlikely the trial will reach a positive conclusion, in favour of the new treatment,

Then, the trial may be stopped "for futility", saving resources for other studies.

My research: Efficient stopping rules for Phase III trials

Starting in my PhD, nearly 40 years ago, I have carried out research into methods for (group) sequential monitoring of clinical trials.

Procedures should protect the type I error rate and guarantee the desired power.

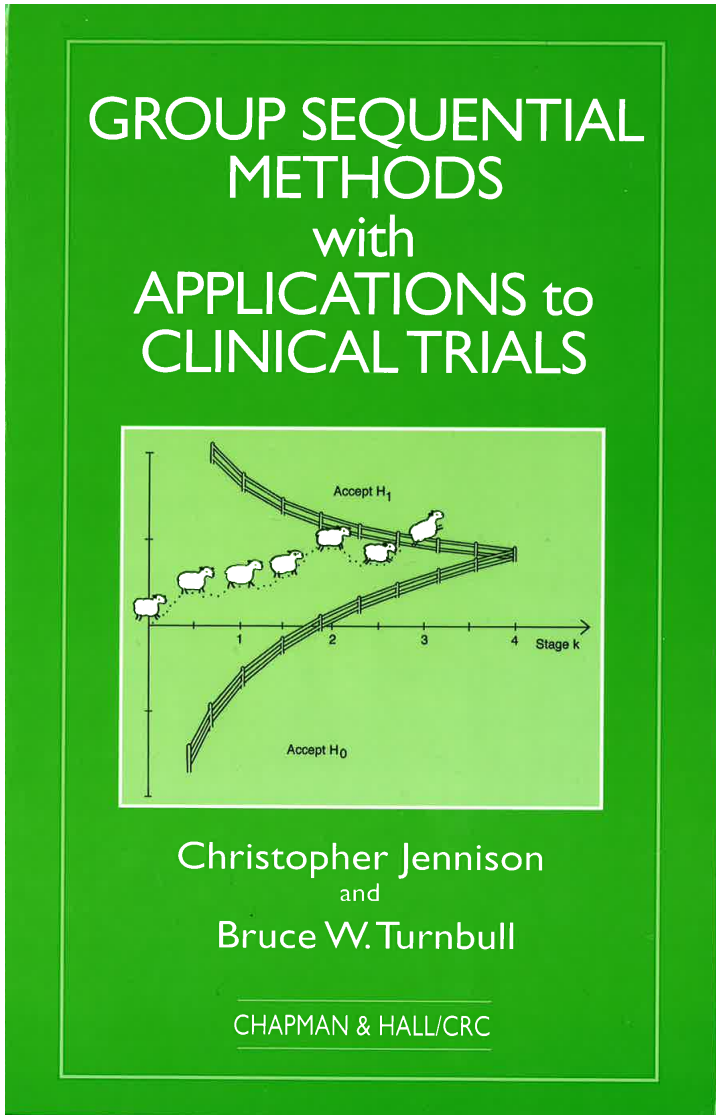
At the same time, the trial should stop as early as possible, minimising the average number of patients or average time to a positive conclusion.

This work involves

- Deriving theory for the distribution of a sequence of estimates of treatment effect, for different types of primary endpoint
- Optimisation of stopping rules
- Defining flexible procedures to deal with unpredictable levels of information at each analysis

Phase III trials: Early stopping

See



Adaptive clinical trials

In a traditional approach, all aspects of a clinical trial are specified at the planning stage, before the first subject is treated.

This helps to ensure results are easily interpreted and there is no suspicion that the trial's goals or methods of analysis have been adjusted to suit the data observed.

On the other hand, many important properties of the treatment, the patients, and outcome variables, are not known before the trial.

Knowledge about these points would have helped in designing the trial. So, can we "learn as we go"?

Around the turn of the millennium, proposals began to appear for "adaptive trial designs", that use information gained during the trial to re-design the trial's objectives and ways to achieve these.

We shall focus on adaptation in Phase III trials.

Updating the sample size

A major goal of adaptation is to update a trial's sample size to ensure it has adequate power.

With a survival or other time-to-event endpoint, changes are made to achieve the desired total number of events.

For a normally distributed endpoint, sample size may be increased if the estimated response variance is higher than the value assumed in the original sample size calculation.

There have been proposals to modify sample size in response to estimates of the treatment effect — but I would recommend a group sequential design that starts with a suitably high maximum sample size and stops early when the data allow this.

More exotic adaptations

Some interesting statistical challenges arise in trials that involve multiple treatments or multiple patient populations and sub-populations.

These include

- Seamless Phase II/Phase III,
- Adaptive enrichment trials,
- Multi-arm Phase III trials.

If done well, these types of trial designs can answer questions more accurately and more efficiently.

A multi-arm Phase III trial

The GATSBY trial: A study of

- Trastuzumab Emtansine vs Taxane
- in patients with Human Epidermal Growth Factor Receptor 2 (HER2)-Positive Advanced Gastric Cancer

sponsored by Hoffmann-La Roche.

The trial compared:

Experimental Treatment 1: Slower dosing,

Experimental Treatment 2: Intensive dosing,

Control treatment.

The GATSBY trial: Details of treatments

Standard Taxane Therapy Docetaxel was administered at 75 mg/m^2 IV on Day 1 of a 21-day cycle, or paclitaxel was administered at 80 mg/m^2 IV weekly (Days 1, 8, and 15 of a 21 day cycle) as per investigator's choice, until progression of disease, intolerable toxicity, initiation of another anticancer therapy, or participants or physician decision to discontinue.

Trastuzumab Emtansine 2.4 mg Trastuzumab emtansine was administered on Days 1, 8, and 15 of a 21-day cycle at 2.4 mg/kg IV infusion until progression of disease, intolerable toxicity, etc.

Trastuzumab Emtansine 3.6 mg Trastuzumab emtansine was administered on Day 1 of a 21-day cycle at 3.6 mg/kg IV infusion until progression of disease, intolerable toxicity, etc.

See *ClinicalTrials.gov*

The GATSBY trial

Initially, patients were randomised to the three treatment arms.

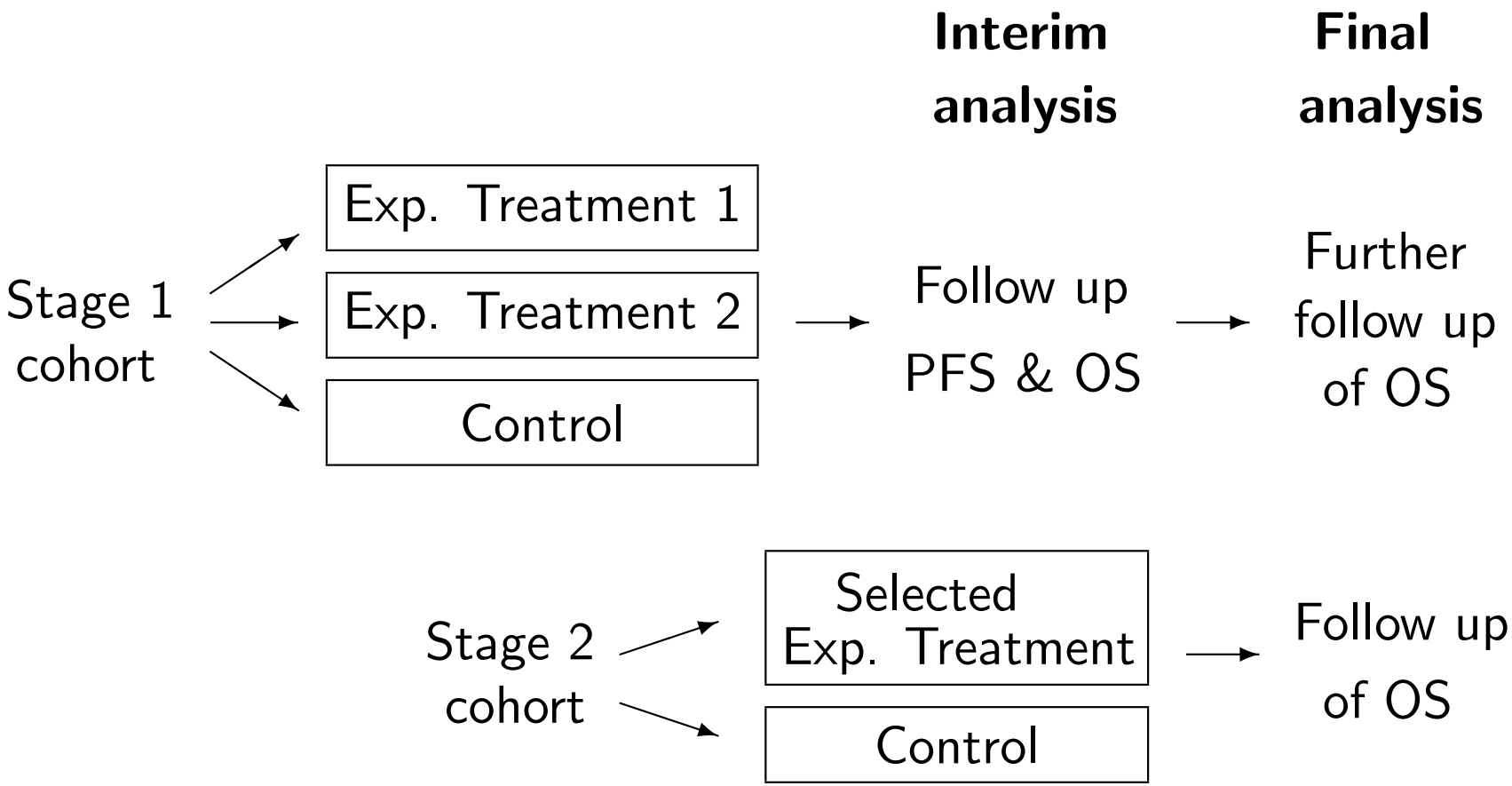
At an interim analysis, one of the two experimental treatments was selected and new patients were randomised between the selected experimental treatment and the control.

The primary endpoint was Overall Survival (OS).

Information on OS, Progression Free Survival (PFS), safety, and PK measurements were used at the interim analysis to choose between the two experimental treatments.

Note that PFS is useful here as it is more rapidly observed.

Overall plan of the GATSBY trial



At the final analysis, we test the null hypothesis that OS on the selected treatment is no better than OS on the control.

The GATSBY trial: Statistical issues

Careful attention is needed to protect the type I error rate in the presence of:

- Multiple testing,
  - Multiple null hypotheses, one for each experimental treatment vs control
  - Data dependent choice of the null hypothesis tested in the final analysis
- Combining data from the two stages, before and after adaptation — particular care is needed for survival data.

For this type of design to be advantageous, there needs to be adequate information from PFS data, etc, to inform the choice between experimental treatments at the interim analysis.

The GATSBY trial: Statistical issues

What are the benefits of designing the trial this way?

It can be shown that, under certain conditions, a trial design with treatment selection at an interim point is more likely to confirm that the experimental treatment is superior, when this is indeed the case.

A subsidiary feature is that the trial allows a head to head comparison of the two types of schedule for administering Trastuzumab Emtansine.

The GATSBY trial: What was the outcome

Participant Flow:

	Standard Taxane Therapy	Trastuzumab Emtansine 2.4 mg	Trastuzumab Emtansine 3.6 mg
STARTED	117	228	70
Stage 1	37	75	70
Stage 2	80	153	0
Death	90	187	61
Lost to Follow-up	3	2	1
Withdrawal by Subject	14	11	5
Study Terminated	10	28	3

The GATSBY trial: What was the outcome?

The trial was conducted according to plan.

The Trastuzumab Emtansine 2.4mg treatment was chosen at the interim analysis.

The trial continued to completion but, in the final analysis, there was no significant difference in overall survival between patients treated with Trastuzumab Emtansine and patients treated with Taxane.

Conclusions

- There are opportunities for innovation in the design of clinical trials — and these can lead to logistical and statistical problems that need to be solved.
- Recent developments in statistical methodology can support a range of new methods.
- We know how to conduct clinical trials group sequentially or adaptively and still maintain control of the type I error rate.
- The challenge of making these trial designs as efficient as possible will keep us busy for the foreseeable future.