Running a Group Sequential Clinical Trial

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

Department of Mathematical Sciences,
University of Bath

PSI Conference, London, 2023

1. GSDs: What, why and how?

A group sequential design (GSD) allows investigators to stop a clinical trial as soon as there is evidence to support that decision.

This may be for

A positive outcome when:

There is clear evidence the new treatment is superior to the control, and

There is adequate safety data, and

No further information on secondary endpoints is required.

A negative outcome when:

The new treatment has an unacceptable level of serious adverse events, or

It is highly unlikely that the trial will show the new treatment to be superior.

GSDs: What, why and how?

How early can a trial stop?

Suppose type I error rate α and power $1 - \beta$ at a treatment effect δ are specified.

With a small number of interim analyses and a maximum sample size 5% or 10% higher than the fixed sample size, a GSD can reduce average sample size to around

70% of the fixed sample size.

In a time-to-event study, a similar reduction can be achieved in the number of observed events.

Potential benefits

- Save costs by recruiting a smaller sample size,
- Submit a New Drug Application sooner,
- Make a new treatment available to patients earlier.

GSDs: What, why and how?

How to design a group sequential trial

Start by creating a fixed sample size design.

Specify the primary endpoint,

Set up a hypothesis testing problem with type I error rate α. and power $1 - \beta$ if the treatment effect, θ, is equal to δ,

Assign values to parameters such as response variance or event rate in order to carry out a sample size calculation.

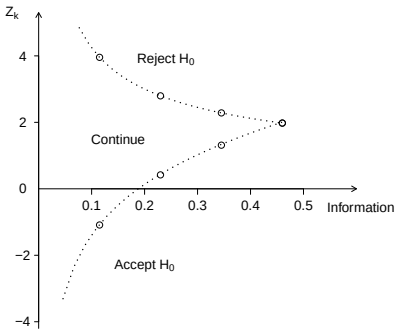
Modify this fixed design to create a group sequential design.

Choose an error spending function (or functions),

Set the number of planned analyses,

Calculate the new (slightly higher) maximum sample size.

Running a trial with an error spending design



At each analysis, calculate the stopping boundary.

Plot the current data point and see whether the trial should terminate or continue.

Error spending designs: Some technical details

In an error spending design, type I and type II error probabilities are "spent" as a function of the observed information, so the stopping boundaries satisfy

$$P_{\theta=0}\{\text{Reject } H_0 \text{ by analysis } k\} = f(\mathcal{I}_k),$$

$$P_{\theta=\delta}\{\text{Accept } H_0 \text{ by analysis } k\} = g(\mathcal{I}_k).$$

Here information is defined as $1/\text{Var}(\widehat{\theta})$, where $\widehat{\theta}$ is the current estimate of the treatment effect, θ.

The maximum sample size, or maximum follow-up for time-to-event data, is set to reach a target information level if the trial continues to the final analysis.

Software is available to carry out the computations needed to create and implement error spending trial designs.

2. GSDs: What are the issues?

Some questions

1. At the design stage, there are important things we do not know, e.g., the response variance, or the control arm hazard rate. What should we do?

2. How do we choose the right sort of stopping rule?

Some answers

1. The same problem arises in designing a fixed sample size trial. “Information monitoring” in an error spending design can be used to adapt sample size or the timing of interim analyses to reach a target information level.

2. The properties of different types of group sequential boundaries are well understood. For a description of some efficient designs see Chapter 7 of *Group Sequential Methods with Applications to Clinical Trials* by Jennison & Turnbull.

3. Group discussion

What are your thoughts?

Are your enthusiastic about group sequential designs or do you have reservations?

If you have conducted a group sequential trial, what issues did you have to address?

If you have considered a group sequential trial and decided against it, what were the difficulties that could not be overcome?

4. Some likely issues

Question 1

Does a GSD have rules or guidelines?

And what happens if the rules are not followed?

Answer:

Stopping rules are usually treated as guidelines and may well be presented as such to a trial's Data Monitoring Committee.

If a trial continues after the efficacy boundary has been crossed, this can only lower the type I error probability.

A GSD can be constructed with a "non-binding futility boundary". Then, the type I error probability is calculated assuming the trial will never stop for futility. Since stopping for futility may occur, the type I error rate is achieved conservatively.

Some likely issues

Question 2

We cannot stop the trial early as we may need more safety data to submit a New Drug Application.

Answer:

Early stopping for futility is still possible (as there will be no NDA).

Is there a point during the trial at which the number of patients on the new treatment is sufficient to assess safety? If so, one may create an efficacy boundary that allows early stopping after this point.

If observing a certain number of patients on the new treatment is a key objective, you may consider randomising a higher proportion of subjects to the new treatment throughout the trial.

Some likely issues

Question 3 Interim analyses are costly: Are more analyses really worthwhile?

Answer: The benefits of group sequential testing can be substantial.

Consider GSDs with binding futility boundaries, K equally sized groups, $\alpha = 0.025$, $1 - \beta = 0.9$ and $\mathcal{I}_{max} = R\mathcal{I}_{fix}$, where $\mathcal{I}_{fix}$ is information in a fixed sample design.

Minimum possible $\{E_0(\mathcal{I}) + E_\delta(\mathcal{I})\}/2$ ***as a percentage of*** $\mathcal{I}_{fix}$

K	1.01	1.05	1.1	1.2	1.3
2	80.8	74.7	**73.2**	73.7	75.8
3	76.2	69.3	66.6	**65.1**	65.2
5	72.2	65.2	62.2	59.8	**59.0**
10	69.2	62.2	59.0	56.3	**55.1**

Some likely issues

Question 4

What do we do with data that arrive after the decision to stop the trial?

Answer:

Having more information should be a good thing!

This can be the case if these data are handled well.

See the paper “Group sequential tests for delayed responses”, Hampson & Jennison, *JRSS, B*, (2013).

Several groups are working to develop such methods further and explain the methods to potential users.

Some likely issues

Question 5 We wish to analyse secondary endpoints.

Answer:

There are methods that allow inference on a secondary endpoint following a group sequential test for the primary endpoint.

These methods may permit continuation of the trial to test a secondary endpoint after a successful outcome on the primary endpoint.

See, for example,

Glimm, Maurer & Bretz, *Statistics in Medicine*, (2010),

Tamhane, Wu & Mehta, *Statistics in Medicine*, (2012).

There are also group sequential designs that test two endpoints: see

Jennison & Turnbull *Biometrics*, (1993).

5. Concluding remarks

Group sequential tests can have significant benefits:

- Saving costs by recruiting a smaller sample size,
- Submitting a New Drug Application sooner,
- Making a new treatment available to patients earlier.

There may be issues in implementing a group sequential design.

However, this is a well studied area and many potential problems have been solved.