

Technical and practical barriers in applying adaptive designs in clinical trials

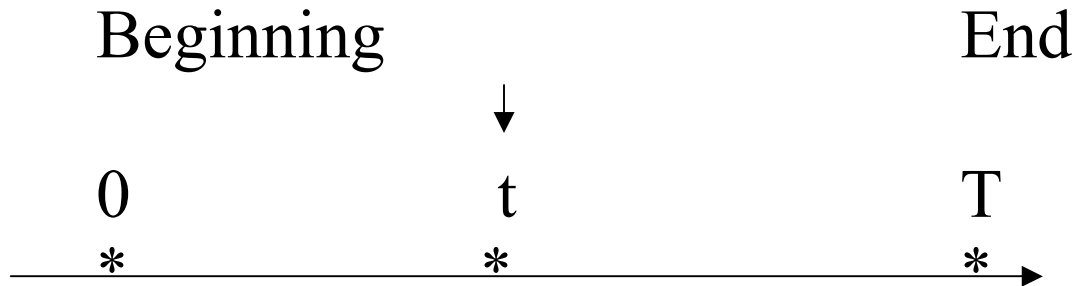
Gordon Lan
Johnson & Johnson
November 13, 2006
Rockville, MD

Outline:

Practical problem: Not enough communication between DMC and the sponsor.

Technical problem: We don't know how to control (or define) α .

Adaptive (Flexible) design



At time 0, protocol was developed.

At time $t > 0$, use the updated “information” to modify the design of the study.

Practical problems:

1. May encourage careless planning of clinical trials.
2. To take full advantage of adaptive designs, interim analysis results have to be communicated, directly or indirectly, to the regulatory agency and the sponsor(s).
At the present time, we do not have a good system to do this without damaging the credibility of the study.

Perception about interim looks, operational bias, inexperienced sponsors. (Lack of trust in the current system.)

If modifying the design of an experiment based on accrued information will improve “efficiency”, then we should create a system to implement this idea.

It is hard to anticipate everything in advance.

Back-up compounds

Lack of toxicity

Financial considerations.....

Increase N from 500 to 650 or 750?

How could we expect the “independent” DMC members always make right decisions for the sponsor?

If serving on the DMC requires taking a “training” short course (2 weeks? one month?).....

	Trial concludes not beneficial	Trial concludes beneficial
Truly not Beneficial (H ₀)	correct	incorrect (Type I) α
Truly beneficial (H _a)	incorrect (Type II) β	correct $1-\beta=\text{power}$

α : Probability of Type I error

β : Probability of Type II error (power = 1- β)

Optimal design (?)

*In statistical research, we would like to propose the use of “an optimal design with an efficient test”. In clinical trials, I have no idea what this term means.

Parameters involved are:

α , $1-\beta$ =power, expense, profit....

Note that “power” depends on treatment effect Δ which is unknown to us.

Design a trial under the constraint of $\alpha=0.025$.

If we treat prior information as “informative” then:

Frequentist: Increase α to a higher level.

Bayesian: Use a favorable prior.

How can we incorporate prior information into the trial data in an objective manner?

Most of the time, we ignore the prior information just because there is no simple way to quantify it in an objective manner.

How can we control the α -level when there are multiple endpoints?

Suppose there are 3 endpoints, (X_1, X_2, X_3) .

Can we find a univariate response index

$U = \delta_1 X_1 + \delta_2 X_2 + \delta_3 X_3$ for treatment evaluation?

We usually “pick” a primary endpoint for the comparison of treatments.

Suppose the corresponding Z-values are Z_1 , Z_2 and Z_3 .
with correlation $\rho \geq 0$ for each pair.

Conclude treatment benefit if $Z_1 \geq 1.96$ and $Z_2, Z_3 > 0$.
(We may also require the treatment effects to exceed
certain desirable values.)

What is the chance to conclude benefit under H_0 ?

Probability of $[Z_1 \geq 1.96 \text{ and } Z_2, Z_3 > 0]$

ρ	Estimated α by simulations
0.0	.0000
0.1	.0091
0.2	.0121
0.3	.0153
0.4	.0183
0.5	.0209
0.6	.0232
0.7	.0245
0.8	.0248
0.9	.0250
1.0	.0250

Survival time (X_1) versus a softer endpoint (X_2):

Choose X_2 as primary, conclude treatment benefit if $Z_2 \geq 1.96$.

What if $Z_1 \geq 1.96$?

$Z_1 \geq 1.96$ provides “more evidence” for benefit than $Z_2 \geq 1.96$. (Suppose this is not a QoL study.)

Personal opinion: Just replace the primary endpoint by X_1 without penalty. (Dr. DeMets)

$P[Z_1 \geq 1.96 \text{ or } Z_2 \geq 1.96] > 0.025.$

Split α into $\alpha_1 + \alpha_2 = 0.005 + 0.02 = 0.025$ and claim for treatment benefit if $Z_1 \geq 2.576$ or/and $Z_2 \geq 2.054.$

This strategy could cause many problems during study design and IDMC interim data analyses.

Re-estimation of sample size

Suppose based on an interim unblinded data analysis, we increase the target sample size from N to $N^*(>N)$.

Reason: The extension will increase the Conditional Power (or Predictive Power) to a higher level.

The traditional un-weighted Z test is NOT standard normal.

The weighted Z two-stage approach

$$Z = \sqrt{w}Z_{\text{I}} + \sqrt{1-w}Z_{\text{II}} \quad (w \text{ is pre-specified})$$

Example:

$$Z = \sqrt{0.4}Z_{[1 \rightarrow 40]} + \sqrt{0.6}Z_{[41 \rightarrow 100]}.$$

Suppose sample size is modified from 100 to 200.

$$Z^{\text{W}} = \sqrt{0.4}Z_{[1 \rightarrow 40]} + \sqrt{0.6}Z_{[41 \rightarrow 200]}.$$

$$Z^{\text{UW}} = \sqrt{0.2}Z_{[1 \rightarrow 40]} + \sqrt{0.8}Z_{[41 \rightarrow 200]}.$$

Advantage of the weighted Z-test

Under the null hypothesis, the weighted Z test is always $N(0,1)$. When there are interim looks, the Brownian motion structure behind the Z-values remains unchanged.

⇒ Classical group sequential boundaries can be applied to interim Z-values with minor modifications.

*Make sure that the clinicians understand the implication of the “weighted Z” approach.

(Mathematical convenience versus clinical judgment.)

Adaptive designs are very dynamic. If we always take a “conservative” approach to control α for every possible modification, then the real α -level could be much smaller than 0.025.

$\alpha = 0.025$ → desired power level

$\alpha < 0.025$ → less than the desired level

$\alpha \ll 0.025$ → much less than the desired level

Major technical problem: We don't know how to control α accurately.

Other technical problems:

1. Need more statistical research works on adaptive design of clinical trials.
2. Hard to develop a software that handles all the possible modifications.
3. Lack of simple interpretations for many Bayesian methods.
(Assumptions and interpretations)