

Monitoring and confidentiality issues for adaptive design trials

Paul Gallo

PhRMA Adaptive Designs Workshop

November 13, 2006



1. Within-trial consistency of outcomes

General issue:

- Current monitoring conventions hold that bias can be introduced by knowledge of interim analysis results (objective trial management; attitudes of participants).
- Adaptive designs may be more prone to bias due to the monitoring processes utilized, and because changes will be implemented based on the interim results.
- In order to have some level of assurance about avoidance of bias, can we examine trial data for any signs of changes or biases that are due to knowledge of interim results?

Recent literature / guidance

Koch (*Biom J*, 2006):

- *“Substantial discrepancies with respect to . . . results obtained may then raise a concern: it might be difficult to interpret the conclusions from the trial if it is suspected that the observed discrepancies are a consequence of (intentional or unintentional) dissemination of the interim results.”*

CHMP draft reflection paper (2006):

- *“Studies with interim analyses where there are marked differences between different study parts or stages, will be difficult to interpret. It may be unclear whether the effects differ just by chance, or as a consequence of the intentional or unintentional communication of interim results, or for other reasons . . . the applicant must pre-plan methods to ensure that results from different stages can be justifiably combined.”*

Questions / issues

What could be a standard for establishing that a change took place?

Evidence of interaction (suggestion of differential treatment effects before and after interim analyses) – what strength of evidence?

Might a *main effect shift* also be of some concern?

Should we necessarily ascribe changes to knowledge of interim results?

Implementing a formal approach

Can we implement some type of standard, formal approach?

- Conventional statistical significance, e.g., from an interaction test, might be too high a standard.
- Looser test standards (e.g., $p=0.10$, 0.20) raise risks of falsely suggesting impaired interpretability of results (impact on '*power*'?).
- If we look for 'signals', we probably can expect to find them frequently.
- Are some changes *natural*, or at least unrelated to the interim analysis? (natural drift in the patient population; 'learning curve'; etc.)

Quiz

We design a 2-arm study to detect a specified effect Δ (2.5% level, 80% power).

We further assume that, in fact, Δ is the true effect.

If we divide the trial data into 3 equal parts, what is the chance that in one of those stages we would see a point estimate that is more than 50% away from Δ ?

- 20%
- 40%
- 60%
- 80%

Quiz (part II) . . .

Same set-up. Now we divide trial data into 2 equal halves, estimate the treatment effect within each (d_1 and d_2)

What observed magnitude of interaction would reach conventional significance?

➤ $|d_2 - d_1| = 1.4 \Delta$

What would be the power for detecting, between the first and second halves of the trial, a true increased treatment effect of magnitude $\Delta/2$?

➤ 10.4%

Practical implementation issues

Formal approaches and standards seem challenging because the operating characteristics seem likely to be poor.

Is the 'cutpoint' clearly defined?

- e.g., data used for the IA, DMC meeting, implementation of adaptation, etc.
- *Patients* vs *outcomes*: patients enrolled before the IA but on study into the next stage (what is the 'lag' before any biasing mechanism would introduce change?)

What factors should be adjusted for in this investigation?

Causality of any changes seen.

What *can* we do?

- Koch (2006): “. . . *confidentiality can hardly ever be proven . . .*”
- Numerical results obviously have a role to play.
- Perhaps this issue is more of a concern in situations where the mechanism for potential bias is more apparent?
- Convincing *procedures* to protect and restrict information, and *documentation* that these were followed, will be ***critical***.

2. Sponsor involvement

Typically, sponsor is uninvolved in data monitoring, for maximum objectivity of decision making and integrity of trial results (e.g., FDA DMC guidance).

In adaptive trials:

- Might sponsor perspective be relevant for most effectively making certain complex types of adaptation decisions?
- Will sponsors accept and trust decisions made confidentially by external DMCs in long-term trials / projects with important commercial implications?

Sponsor concern: *unanticipated complexities* that might not fit a pre-specified algorithm.

Proposed model for sponsor involvement

(Let's envision a long-term seamless phase IIb/III program as a model for this discussion.)

Potential sponsor participation in the process in confirmatory adaptive trials should require:

- a clear rationale for the involvement, based on the complex nature of the decision and its implications
- individuals properly '*distanced*' from trial operations
- clear understanding by all involved of the issues involved and risks to the trial, documentation of the processes followed, and restrictive firewalls / procedures in place

Model for sponsor involvement (*cont.*)

- “minimal” sponsor exposure to make the needed decision:
 - *smallest number of individuals* who can supply the needed perspectives
 - *only at the adaptation point* (which often will be at a relatively early point within the trial)
 - *only the minimally relevant information* (e.g., unlike a DMC with whom they may interact, which may have a broader ongoing role).

(Note: see Section 6.5, FDA-DMC Guidance !!)

Sponsor involvement

The burden will necessarily fall on the sponsor to make a strong case that effective procedures and safeguards are in place, and are followed.

Restricting information should be a small price to pay for the advantages that these designs offer.

Balancing *trial integrity* versus *sponsor interests* ??

➤ *Trial integrity is a sponsor interest !!*

3. Information conveyed to observers

- Adaptive designs have been questioned on the basis that changes made are *apparent* to some extent, and might thus provide some information to observers about the results which led to those changes.
- Understanding the principles which are the basis for the confidentiality concerns: can't we distinguish between *types* and *amounts* of information, and how problematic they might be in this regard?
- Example: even some conventional group sequential designs convey more information than seems to be commonly perceived.

Information conveyed to observers

There would seem to be a difference, for example, between:

- changes made in an algorithmic manner based on an interim treatment effect estimate, which can allow an observer to back-calculate that estimate (e.g., certain sample size re-estimation approaches)
- selection decisions, such as might be made in a seamless phase II/III design (e.g., which dose group is selected to continue).

Information conveyed to observers

Proposal

Selection decisions, of the type made in seamless designs (e.g., choice of dose, subgroup, etc. for continuation), and with the specific numerical results on which the decisions were based remaining strictly confidential, do not in most cases seem to convey an amount of information that should influence or compromise a trial.

For the panel . . .

1. How can we address in trial data whether there is evidence of inconsistency of outcomes across stages, suggestive of bias due to knowledge of interim results?
2. Can there be a properly-restricted role for sponsors in certain types of adaptation decision making?
3. Information conveyed by observing adaptations: are selection decisions OK?