

# Myth Busting - Clinical

Robert J. Temple, M.D.

Associate Director for Medical Policy  
Center for Drug Evaluation and Research  
U.S. Food and Drug Administration

PhRMA

Adaptive Design Workshop

November 13, 2006

# Adaptive Designs

- Not a wholly new thing.
- Not a substitute for an effective drug, in the right population, at the right dose (but perhaps a way of finding those things).
- Still need to resolve all the usual worries about clinical trials, such as inflation of alpha error (multiple analyses, subgroups), introduction of bias from examination of interim results and developing data-conditioned hypotheses. It is not a way to look at data and do anything you want to.
- Sometimes pauses for analysis lead to insight and intelligent modification. These may not be possible in “seamless” approaches and could undermine advantages. Experience will tell us (if we look).

# Adaptation up to Now

Adaptive designs, most broadly, could include any design where a critical aspect of the study is altered based on results.

Some are routine (or at least not rare)

## A. Blinded sample size adjustments

- Adjusting sample size based on number of endpoints (blinded) in outcome trials. We urge studies NOT to specify sample size but endpoint numbers. No statistical correction.
- Adjusting sample size based on variance at interim look, now recommended for pediatric written requests for antihypertensives to assure adequate power, but potentially much more broadly applicable. Again, no correction.

# Adaptation Up to Now (cont)

## B. Result-dependent analyses

It is possible to plan alternative pathways for analysis, handling drop-outs, etc. It is common to specify covariate analysis if \_\_\_\_\_, \_\_\_\_\_, \_\_\_\_\_ and even to say what would get a particular covariate into the analysis, and to specify different analysis (Cox vs M-H) depending on results. As these are pre-specified adaptations, no correction is needed.

# Adaptation Up to Now (cont)

## C. Starting extra dosing groups

To study possible toxicity of the higher dose. Used for iib/iii inhibitor studies.

PURSUIT: eptifibatide vs placebo in 10,000+ patients with ACS using a bolus of 180  $\mu\text{g}/\text{kg}$  and two infusions (1.3  $\mu\text{g}/\text{kg}/\text{min}$  or 2.0  $\mu\text{g}/\text{kg}/\text{min}$ ). After 3218 patients, with no XS bleeding, low dose was dropped.

No statistical correction because early endpoint and late endpoint are totally distinct.

# Adaptation Up to Now (cont)

## D. Group-sequential analyses

The norm in outcome studies and considered an ethical imperative. The interim endpoint and final primary endpoint don't have to be the same.

Because with early stopping you lose data (duration, possible subset effects) and because some endpoints are more secure, where the study has a combined primary endpoint (e.g., death plus heart failure hospitalization) we have urged stopping for mortality but not for other components of the combined endpoint.

An interesting issue is what sort of correction is needed for the primary (combined endpoint) if you would stop only for mortality, which is only a fraction, usually well under half, of the primary endpoints.

# Adaptation Up to Now (cont)

E. Enrichments – these raise no statistical issue and are “pre-study adaptations” that can add enormously to efficiency

1. Screening (sort of adaptive)

- Pharmacodynamic or clinical response before randomization
  - Sublingual NTG for entry into topical NTG trials
  - VPB suppression in CAST before randomization
  - Could be the clinical response being measured
- Single blind placebo lead-in periods prior to randomization
- Tolerability of BB before randomization in CHF trials
- Tumor response single arm: treat responders; drop progressors; randomize stable disease (Simon)
- Can imagine using imaging response to choose people to randomize in cancer, Alzheimer’s Disease, stroke studies, perhaps even anti-depressant trials

# Adaptation Up to Now (cont)

## 2. Late assessment of baseline characteristic

- Evaluate only patients with sensitive organisms in antibiotic trials
- Late assessment of a tumor marker and prospective analysis of the (+) subset
- Other cases where randomization is urgent but interest is only in a later identified subset, planned prospectively

# Adaptation Up to Now (cont)

## II. Not Routine But Contemplated or Possible

### A. Unblinded sample size adjustments.

Power estimates are always guesses. Methods have been described for enlarging studies based on interim unblinded analyses.

### B. Using pharmacologic endpoints/biomarkers to modify study.

Include several dosing groups and let early tumor response decide which doses to carry forward. As response and long-term outcome could correlate, need some correction but if only a small fraction of total sample is used, could be very minor.

# Adaptation Up to Now (cont)

## c. Play the winner designs

Rarely seen, not clearly more efficient, but we have not objected in concept.

## D. Minimization designs

Rarely seen, ongoing debate about whether randomized.

# Adaptation Up to Now (cont)

E. Using part 1 to modify part 2

Friedlin and Simon proposed carrying out half a study then looking at results for a response predictor (genetic, other).

The study would then be completed with two primary endpoints:

- Whole study at  $p = 0.04$
- Responsive subset, 2<sup>nd</sup> half only, at  $p = 0.01$

Would work only for very large subset effect

but would, e.g., for erlotinib (Tarceva,

Herceptin) where response in “off group” is very small and in identified responder subgroup is large.

# Adaptation Up to Now

## F. Two farther out possibilities

1. Randomized WD in a subset similar to a clinical screen (randomized only the responders), could do a randomized withdrawal study, even in an overall failed study, in a responder subset.

I've never seen this.

2. Retrospective replicated subset analysis

Search a 50% sample for a responder subset and verify in the blinded remainder. Could be one A & WC study.

Doing in thirds would be more secure.

Needs discussion.

# Finding the Population and Dose

The adaptive dream is to use early study data to redirect critical aspects of the trial, to

- Right dose
- Right subset
- Right endpoint
- Right time of analysis

But the more opportunities to choose from, the greater the increase in alpha error, so, where possible, choosing before the study may ultimately be more efficient. That seems quite possible for

- D/R for a biomarker
- Randomizing only screened responders as opposed to analyzing a subset
- Getting a reasonable dose range to study (modeling/biomarkers), Sheiner-Beal nonmem
- Anticipating outcomes and consequences of analysis choice and pre-specifying

# Making Sure Changes Meant to Be Blinded Are Blind

We are seeing increasing number of trials with changes in SAP at or beyond study completion date. We ask for evidence of blinding, but it's a major problem. Critical to fully define and describe "firewalls" in place and assure unbiased changes.

# Consider Downsides

## 1. Pauses can refresh

The phases were not designed as punishment but as a description of usual thinking about development. Expected that phase 2 would involve learning (Sheiner) and modification of many aspects of trials (dose, interval, endpoints, concomitant Rx, titration for tolerability, endpoints).

If it's all one trial, you lose those chances.

We already have inadequate dose finding, reluctance to study multiple doses in phase 3, with consequences sometimes (Lotronex, astemizole). Could “seamless” designs make it worse?

# Consider Downsides

## 2. Are there alternatives

“Seamless” trials avoid IRB and internal delays. Might an IRB or internal monitor consider a pair of trials that would be fully designed except for information from a phase 2 trial that would inform dose, dose interval, monitoring, so that, barring significant change, abbreviated review could suffice.

As noted, selection via screening avoids many problems of adaptation.