

The Use of Adaptive Designs in an NIH-Funded Clinical Trial Environment: A Discussion of the Presentations

J.L.P. (Seamus) Thompson, Ph.D.

Professor of Clinical Biostatistics
and Clinical Neurology
Columbia University

Society for Clinical Trials Annual Meeting
Atlanta GA, USA
May 6, 2009

TOPICS INCLUDE

“issues with budget administration, the timely collection of quality data, and the manner in which decisions (possibly involving the sponsor) are made and implemented in ways that ensure that the credibility and integrity of the trial are not compromised”

- Professors Palesch and Chaloner address Phase I (dose selection), and Professor Levin addresses phase II and its interface with phase III.
- There is a concentration in neurology (NINDS), but the issues are general across agencies. Professor Chaloner's work is sponsored by NIAID.

1 PALESCH - DFO

- Thanks for an admirably clear presentation of the pros and cons of traditional phase I and CRM designs.
- The DFO experience: initial dialog with clinical PI started after out-of-range initial score, given lack of stat input – familiar.
- If/as Adaptive Designs become more widely used, we can expect the need for statistical consultation to increase.

Reviewer reaction

- NINDS: reviewer enthusiasm for CRM – no doubt helped by Palesch and co's clarity.
- FDA: a different reaction. CRM should be familiar, since it is used in cancer trials, whether or not statisticians are involved.
- CRM's major drawback is that, given binary outcome, it tends to recommend doses above what should be the MTD. But FDA reviewers seem unaware of this; their response focused on possible false recommendation of too **low** an MTD, where CRM performs quite well.

FDA is behind. How will they react to the next generation of designs, e.g.:

CRM with multiple toxicity constraints. This takes account of the probability of grade 4 toxicity: it differentiates between grades of toxicity rather than being limited to a binary outcome (S. Lee, at this conference).

Some holdup, but ..

- Mar 2007 2.4 percentile!
- April 2007 FDA clinical hold: letter August 2007 (4 months)
- Letter lifts hold Nov 2007 – 7 months delay: NGA Jan 2008.
- Trial had not started – award money was not consumed

- Pointless hold before trial start is bad enough
- But suppose there was a long inappropriate hold after trial start, training etc,
- and grant funds were consumed without enrolling patients
- and performance was evaluated by the NIH on the basis of recruitment and followup during those months? – see TNK.

2 CHALONER - H-Flu

- Modified Up/down design, nicely described.
- 2 stages with 15 patients in all
- 6+9: 1st stage 1 pt at a time, 2nd 3 at a time

- 80% probability of no MLE estimate stage 1, 40% stage 2.
- MLE apparently obtained – seems quite close to Bayesian estimate. A Bayesian analysis with good frequentist properties?
- Sample size too small to assess significance of difference between MLE and Bayesian results.
- When/if you can not get the MLE estimate, what do you have with the Bayesian, and how do you evaluate it?

Reviewer response

- NIH program officers “puzzled”, but reassured
- How deep did the discussions with them go? If not far, we may expect trouble later. If yes, perhaps characterize. [Clarified: discussions were quite extensive.]
- Intuitive appeal to clinicians does not ensure statistical adequacy.
- FDA also had some concerns: again alleviated after conference call meeting.
- Discuss/contrast w DFO experience?

3 LEVIN - QALS and TNK

- Major innovations in QALS (phase II) and TNK (phase II + III). Well described. Agency gets credit for support for innovation.
- QALS worked – CoQ10 futile. Ideally suited to where effective therapies do exist (tPA), but are rare - Neuroprotectants for stroke 0 successes in 20.
- But TNK-S2B, with sound design, and prospect of improving on tPA safety profile, was stopped during phase II.
- Recruitment was a major problem. But not of greater order than other continuing trials?
- Major compounding factors were FDA review delay in face of well-presented analyses; ad hoc reviewers not familiar with the adaptive design; and administrative arrangements designed for fixed budgets, not variable length phase II trials.
- Result: promise unfulfilled. The task is to avoid this being the wave of the future.

CONCLUSIONS

Areas of need and difficulty

- Definition of adaptive designs
- Data management resources
- Funding and budget flexibility
- Reviewer expertise - NIH and FDA
- Do we expect more problems at some stages than others?
- Series of adaptive trials? Pros and cons.

1. Definition of “Adaptive Design”

- We agreed not to get bogged down in this, and succeeded.
- Industry definition is somewhat imprecise – adaptive designs modify “aspects” of designs during trial. But “aspects” are undefined, and have now expanded to a very wide range of options.
- If we adopt a broad definition of adaptive designs such that any alpha-spending or futility analysis is an adaptive design, etc, then we must agree we are seeing a quantum leap upwards in the range and scope of adaptive design options.

- The question is: how well are investigators, grant applicants, and sponsoring agencies equipped to navigate this undoubtedly unfamiliar territory?
- The distinction between prespecified and reactive/ad hoc adaptive designs is critical. Endorse the consensus that the priority is the former, since we cannot assess and defend the latter statistically (Coffey, and many others)?
- There may ultimately be some paradox here. Outside some range, isn't "prospective [specification of] adaptation" something of an oxymoron or impossibility?
- There may be limits, particularly given the call for simplicity, and so many options.

2. Data Management

- High data quality and availability are critical (see e.g. Chaloner on small Ns, + dose delivery problem)
- We can foresee problems if adaptive designs are implemented by investigators not adequately resourced for data management.
- Palesch unit has resources. See also
R Buchsbaum, P Kaufmann, A Barsdorf, R Arbing, J Montes, and JLP Thompson, for the QALS Study Group. Web-based data management for a phase II clinical trial in ALS. Amyotroph Lateral Scler 2008 (DOI 10.1080/17482960802378998
- A caveat: Not all web-based data management systems are created equal.....

3. Funding and budget flexibility

- The NIH model seems to require provision of funding through phase III of any phase II that is initiated.
- Admirable for fixed-length phase IIs with high expectation of success.
- But disastrous for futility-type phase IIs designed to rule out a high proportion of candidates. Most of these phase IIIs will NOT be conducted. This would be success, saving millions of scarce \$\$\$\$. A great potential payoff for adaptive designs is thus threatened. This is a critical problem.
- Similarly, the system not well equipped to fund phase IIs of variable length, as Levin described clearly. Another critical problem.

4. Reviewer expertise

- Always an issue with innovations: well documented here.
- Qualified reviewers are needed at FDA and NIH; and also on DSMBs and IRBs.
- If many new types of adaptive designs are implemented simultaneously, there will be difficulties.
- Some narrowing of options seems appropriate
- An upcoming conference, November 2007, for FDA, non-profits, NIH, and academics may consider this.

5. Expect more problems at some stages than others?

- Yes, at the phase II/III interface, for reasons discussed.
- Phase I trials also have issues, but they are smaller, and of shorter length, and do not pose such severe funding issues in terms of funding allocation.

The Bottom Line

- Adaptive designs hold major promise, but major issues must be addressed if this is to be realized.
- Thanks for the timely session and well-focused presentations!