Reviewer's report

Title: Flexible trial design in practice: stopping arms for lack-of-benefit and adding research arms mid-trial in STAMPEDE: a multi arm multi stage randomised controlled trial

Version: 1 Date: 9 July 2012

Reviewer: Allan Donner

Reviewer's report:

Major comments

This well-written paper provides considerable operational detail and background information to an adaptive trial which is characterized by an innovative multiple arm/multiple stage design.

1. In section 3.2 the terms “power” and “significance level” are introduced without any further discussion in the paper until table 1. However this table merely specifies some input values for these parameters as well as for a “target hazard ratio”. I do note that a paragraph later there is reference to the Stata programmes nstage and art pep that were used to obtain the number of control arm events required to trigger the end of each stage and to estimate the likely duration of the trial, respectively. However, aside from the fact that I would generally prefer to see statistical procedures described or referenced rather than software programs only, it is not clear what methodology was actually used to determine the sample sizes for the trial.

2. The approach taken to the data analysis is also only peripherally alluded to, although the reference to a hazard ratio implies the analysis primarily consisted of a logrank test or Cox proportional hazards regression. If the latter, how was the issue of covariate adjustment handled?

3. A trial of this complexity (up to six arms, multiple stages) must inevitably raise challenges in how to administer informed consent to the participants. Readers would benefit from some description on how these challenges were met.

4. The complexity of this trial also inevitably raises issues concerning multiple comparisons, particularly relevant in the context of those pairwise comparisons which are to be regarded as confirmatory. The well-known method proposed by Dunnett is often used to control type 1 error in trials with this structure, although there is no indication it was applied here. Some discussion on this point would be helpful.

If the issues raised above have been discussed in detail in other publications, this should be explicitly indicated in the appropriate sections of this paper.

Level of interest: An article of importance in its field
Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:

'I declare that I have no competing interests'