Reviewer's report

**Title:** The Healthy Eating and Lifestyle Programme (HELP) study protocol: an efficacy randomised controlled trial of the HELP intervention compared with enhanced standard care of obese adolescent in the community

**Version:** 3 **Date:** 1 August 2011

**Reviewer:** Andrew Vickers

**Reviewer's report:**

1. The randomization section is inadequate (indeed, it appears to end with a sentence that is incomplete). The authors must specify the exact logistic steps taken to ensure full allocation concealment. I would also recommend stratifying on baseline BMI.

2. There are a very large number of secondary outcomes: I counted 28. There are several problems.
   1. How will all this information be interpreted? For example, what if social on the PedsQL is positive but the IWQOL negative? I felt that the secondary outcomes were thrown together because they seemed sort of sensible at first blush, not because it had been rationally assessed why they would add pertinent information.
   2. When will the measures be taken. This wasn't always clear. I suggest a table with columns as time points and rows as outcomes with X’s marking assessment times.
   3. Will the secondary outcomes be analyzed by ANCOVA approaches (i.e. With baseline as covariate)? This isn't clear.

3. I can't agree with the CACE analysis. Compliance is the whole deal here! In the case of a drug therapy, you might say "let's do a CACE to get at the biological effect of the drug, adjusting out for non-compliance." This doesn't make sense for a weight loss programme of counselling and advice.

4. The sample size calculation is totally wrong! In brief, because the main analysis is ANCOVA (which is entirely correct), the sample size calculation has to take into account the correlation between baseline and follow-up measures (see Frison and Pocock Stat Med 1992). My guess is that the correlation between baseline and followup BMI will be very large, drastically reducing sample size requirements. On the other hand, the authors are looking for a very large effect size. My guess is that, with a correlation between baseline and follow up BMI of 0.85, the investigators would have 80% power to detect a d of 0.3