Reviewer's report

Title: Protocol for a randomized controlled study of Iyengar yoga for youth with irritable bowel syndrome

Version: 3 Date: 19 November 2010

Reviewer: Andrew Vickers

Reviewer's report:

1. The randomization section is insufficiently described. The authors merely state that “participants will be randomly assigned to the IY program or the wait-list control group”. Sufficient detail must be give to ensure that allocation cannot be guessed before a patient is unambiguously registered nor changed afterwards. See http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2596474/?tool=pubmed

2. The sample calculation is simply wrong. First, the correlation between baseline and follow-up measures is given as 0.5, which is reasonable, but the authors do not appear to go on and then use this to adjust standard deviation (see http://www.ncbi.nlm.nih.gov/pubmed/1485053, or http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1310536/?tool=pubmed for a simple introduction). Then the calculation itself is off: the authors state that the calculation gives an estimate of 26, and so slight more that 26 will be accrued per group. But their assumptions give 26 patients in total. This is because their effect size is enormous: 0.5 is normally seen as pretty big, the authors assume an effect size well over twice as large.

3. The statistical analysis plan is way too complex. Considerations of distributions of data are misplaced (see http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1310536/?tool=pubmed) and ANCOVA is preferable to ANOVA (see http://www.psychosomaticmedicine.org/cgi/content/full/67/4/652).

4. It is claimed (page 27) that imputation will be conducted for missing data. But this is very complex – certainly difficult to do in SPSS – and would need to be described in much more detail (or removed).

5. The authors state that “The first set of analyses will compare the initial yoga and initial control group”: “Several potential covariates will be examined prior to final statistical modeling to determine whether there are pre-existing differences between the groups. These exploratory analyses will test for differences in demographic, clinical and psychological aspects (e.g., depression). Any variables found to vary significantly by group will be included as covariates."Comparisons of groups at baseline is irrational because it is testing a hypothesis that is known to be true. The data dependent approach to variable selection is also unsound: if it is important, adjust for it, because it will help precision; if not, don’t worry about it. This is a randomized trial and is unconditionally unbiased irrespective of baseline imbalance.
6. If I counted correctly, there are more outcome measures than patients per group. There are 11 primary outcome measures alone (IBS, IBS pain, global improvement, 8 domains of SF36). There are 14 secondary endpoints and several laboratory measures. Is this appropriate? What is the rationale for this extraordinary burden of reporting?