Reviewer's report

Title: Subgroup effects despite homogeneous heterogeneity test results

Version: 1 Date: 23 February 2010

Reviewer: Charles Green

Reviewer's report:

Major Compulsory Revisions

1) The authors have demonstrated a useful approach to graphically assessing dimensions which might characterize heterogeneity in a meta-analytic context, however the rationale for doing so requires some attention. The initial presentation of the issue is framed in terms of investigating so-called "clinical" heterogeneity in the absence of evidence for statistical heterogeneity. Juxtaposing statistical v. clinical significance in this case is problematic. Specifically, the criteria for what constitutes clinical significance is relatively vague. In the context of a conventional, Frequentist meta-analysis characterizing heterogeneity as clinically meaningful in the absence of quantitative evidence suggests an unacceptable level of subjectivity. To frame things differently, how would a researcher go about communicating the significance of the "clinical" finding? On the other hand, a second argument for the approach is actually much more compelling: that the graph provides a means for characterizing heterogeneity along salient predictive dimensions. The current reviewer recommends discarding the argument regarding statistical significance and instead, regardless of the Frequentist statistical findings, using the technique to investigate potential sources of heterogeneity. Identification of subgroups may have important implications for clinical applications. Also, as the literature on Bayesian subgroup analysis points out, Frequentist power for detecting the corresponding interactions is generally abysmal. The current method then would provide an excellent means of exploring and displaying potential subgroup effects, particularly in the context of a Bayesian quantification of subgroup effects.

Minor Essential Revisions

Abstract: "IPD" is first used, but did not seem to be defined (although this reviewer may have missed it).

p. 4: Could the reviewer clarify the statement that metanalytic heterogeneity is a function of the effect size measure?

p. 4: "inevitably" Should this be "inevitable"?

p. 5: Fixed or random effects model may be used to model heterogeneity as a function of predictor variables (i.e. model subgroups).

p. 6: "might hard to interpret" Should this be "might be hard to interpret"?

Discretionary Revisions
p.3: "One of the main issues probably concerns heterogeneity..." It might be best to delete "probably".
p.4: "Nowadays" may be a bit informal.

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Needs some language corrections before being published

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

Declaration of competing interests:

I declare that I have no competing interests.