Reviewer's report

Title: Gestational diabetes and pregnancy outcomes - a systematic review of the World Health Organization (WHO) and the International Association of Diabetes in Pregnancy Study Groups (IADPG) diagnostic criteria

Version: 1 Date: 6 February 2012

Reviewer: Robert Lindsay

Reviewer's report:

Wendland et al have provided a systematic review of pregnancy outcomes using either the IADPSG or WHO criteria. This is an interesting analysis and the question whether one or other criterion might be more effective in predicting outcomes an important one.

In general the manuscript is clearly presented.

1) in table 1 it would be helpful to include a column indicating the proportion of women excluded due to treatment. Clearly effect sizes may be smaller where a larger proportion of the most affected women have been excluded and this will be different between studies.

2) for the meta analyses (Figs 2-6 and results) the most notable result is the relative homogeneity of studies using WHO criteria and heterogeneity of results using the IADPSG criteria. This is an important observation that the authors touch on. It would appear to render the estimates using the IADPSG criteria questionable as suggested by the authors (discussion). I wonder if it is reasonable to combine these estimates. It would be worth exploring further- for LGA the outlier study (while conceding that there are only three studies) is the EBDG. The authors mention in the discussion differences between HAPO and the other studies in terms of protocol. A fuller discussion of this would be appropriate- is incomplete fasting the reason for this heterogeneity between studies? It is noticeable that there would appear to be a far larger increase proportionately in cases when the IADPSG rather than WHO criteria are applied in EBDG (7% increases to 19.7%) as opposed to HAPO (11.3% increases to 16.1%). One interpretation of this is that there are a large number of diagnoses using non fasting samples in EBDG. Taken together this casts some doubts as to whether this meta analysis can be carried out. There are few studies with the IADPSG criteria and results look quite different between them- but this of itself is an interesting observation. Another major difference is the blinding in HAPO but not the other studies- this might also serve to diminish effect sizes in those studies and might be discussed.

3) The majority of the data come from the HAPO study. While analysis of the pooled effects of all of the available studies are of interest more emphasis might be placed on the direct comparisons of results using the different criteria within the HAPO and EBDG studies. This overcomes heterogeneity leading from
differing characteristics (eg BMI and underlying genetic risk) between study populations.

Minor issues
Abstract methods missing “s” in random effects
Abstract results preeclampsia misspelled
Methods “laboratorial”

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:
no competing interests