Author's response to reviews

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 (IADPSG) diagnostic criteria

Authors:

   Eliana M Wendland (elianawend@gmail.com)
   Maria R Torloni (ginecologia@terra.com.br)
   Maicon Falavigna (mfalavigna@yahoo.com.br)
   Janet Trujillo (njtrujill@gmail.com)
   Maria A Dode (malicedode@terra.com.br)
   Maria A Campos (ameliaac@terra.com.br)
   Bruce B Duncan (bbduncan@ufrgs.br)
   Maria I Schmidt (maria.schmidt@ufrgs.br)

Version: 2 Date: 7 March 2012

Author's response to reviews: see over
Gestational diabetes and pregnancy outcomes - a systematic review of the World Health Organization (WHO) and the International Association of Diabetes in Pregnancy Study Groups (IADPSG) diagnostic criteria

Dear Dr Jane Norman
BMC Pregnancy and Childbirth Editor

Thank you for the review of our manuscript. You and the referees have made valuable comments and suggestions. We have addressed all the points and we are now re-submitting the manuscript for your consideration.

Editor's comment:

The introduction (first two paragraphs) is confusing - there are issues about (i) glucose load and (ii) cut of values of glucose to determine glucose intolerance. these should be described clearly and separately.

Please mention earlier that this review includes only data on "untreated" participants.

Response
Thank you for pointing this out. We are now stating that we included only untreated women in the Background (last paragraph) and in the Abstract (Methods

The second sentence in the abstract conclusion is unclear - please rewrite.

Response
We have now rephrased it into two sentences to clarify that 1-associations are of similar magnitudes for both criteria, and 2- high inconsistency was seen for those related to the IADSPG criteria. (abstract conclusion)

Reviewer: Shareen Forbes

Thank you for your thorough review. We have addressed all of your suggestions and believe that they improved the clarity of the manuscript.

1. in the abstract it would be better to state that two criteria exist. Rather than the statement “.....are being recommended.” The WHO criteria were never based on pregnancy outcomes whereas the IADPSG are based importantly on pregnancy outcome data. The authors could be more clear re this.

2. in the results in the abstract the authors describe pregnancy related hypertension with respect to the WHO criteria and pre-eclampsia with respect to the IADPSG. Is this correct? Definitions are not described in the text.

Response
Thanks for pointing these issues out. We have made the corresponding (1 and 2) changes in the sections **Background and Results in the Abstract**.

3. **in abstract again to what extent do the authors feel that the performance of the IADPSG criteria are reduced in the context of reducing the numbers examined by over 50% and hence by reducing the power of the study?** in the main text and table 1 it is clear that when HAPO is excluded 21513 subjects remain.

Response
What we meant by performance in the abstract was size of the association, not its precision. We have now changed the **Results section in the Abstract** to make this clearer, presenting this issue in terms of magnitude of the association.

Regarding the small changes in magnitudes of associations seen (1 and 13%) after the exclusion of the HAPO or the EBDG studies, mentioned in the **2nd paragraph in the Discussion**, we believe that the slightly lower associations for the IADPSG criteria after the exclusion of HAPO, are within what could be expected from applying a criteria developed in one setting to another.

3. **in the background 4th paragraph 2nd line “glucose intolerance” should be changed to impaired glucose tolerance. The fasting definition for the WHO for diagnosis of GDM should also be noted (>/=7mmol/l)**

Response
Done.

4. **I wondered whether the inconsistencies viewed across the studies are secondary to the very different ethnic groups included. It is well known that there are differences in birth weights according to ethnicity. Are the authors able to comment for example on the results in this context and do they change in this context. If they cannot report on this they should discuss that this may contribute to the inconsistencies between studies which I note include studies from Pakistan and Japan.**

Response
We agree that ethnicity may explain heterogeneity in LGA associations as well as in preeclampsia. We have now included a brief mention of this in the **Discussion (4th paragraph)**, highlighting a new study recently published by HAPO (Sacks D, Diabetes Care, March 2012).

One hypothesis is that the IADPSG is vulnerable to heterogeneity as it allows diagnosis on the basis of only one out of three possible measures (fasting, 1h and 2h). This may stem from population variability in terms of the probability of being positive by fasting and post load values, as well as in terms of the possibility of having incomplete fasting (drank coffee or tea with sugar, for example).

5. **the authors point out quite rightly that it is too early still to fully evaluate IADPSG diagnostic criteria. This point is well taken.**
6. discussion. 2nd paragraph – better to state that the IADPSG have lower glycaemic thresholds rather than state that it is a milder form of GDM.

Response
We now state: “First, both GDM criteria, but especially the IADPSG criteria, identify lesser degrees of hyperglycemia when compared to other criteria, such as those previously recommended by the ADA.” (Current 3rd paragraph, Discussion)
As the thresholds did not vary that much, the main issue here is allowing positivity of the basis of one of several options.

Reviewer: Robert Lindsay
Thank you for your valuable comments. We considered all of them.

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.

To estimate the correct proportion of women treated is not possible from the available information. Some studies had already performed the exclusion by study design and did not report the proportion of such women. Thus, in the original version sent (Table 1, last column) we indicated which criteria each study used to classify unquestionable hyperglycemia judged at the time and place of the study to require treatment in pregnancy. We can estimate from the literature that these proportions vary between less than 1% (for those studies using the criteria for DM outside of pregnancy) to values between 2-7% (for those using one of the various GDM criteria).

For some studies, all excluded/not included women were treated, for other studies, women were excluded on the basis of having hyperglycemia requiring treatment (sometimes GDM, sometimes reaching DM level of hyperglycemia). Exclusion of these women left us with varying levels of hyperglycemia to be analyzed against outcomes across studies, which is a possible cause of heterogeneity in the associations and also of the small magnitude of the associations in general. We recognized this problem in our original manuscript, but we have now rewritten the third paragraph in Discussion section to make this clearer. We do not think that these exclusions explain most of the heterogeneity found, as they cannot explain why there is greater heterogeneity in the IADPSG criteria analyses.

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.

Response
We recognized the heterogeneity in the meta-analysis for the IADPSG criteria and opted to make a thorough presentation of the whole picture, permitting the reader to evaluate associations from the aggregate, from the subgroups of studies and for the individual studies. In the conclusion, we make clear the need of additional studies on the application of the IADPSG criteria to other settings.

With respect to the causes of the heterogeneity seen for the IADPSG criteria, although incomplete fasting is a general possibility that needs to be considered for all studies (for example, women drinking coffee or tea with sugar and still considering themselves to be in a fasting state), data to evaluate this is lacking. In the EBDG study, women who admitted not being in a fasting state were rescheduled. This is important because the IADPSG criteria, different from any previous criteria, allows the possibility of making a diagnosis of GDM based on a single and quite low FPG. Another possibility for this heterogeneity, for which we now have supporting data (included in the Discussion), is the large heterogeneity within HAPO, with respect to the percentage of GDM women diagnosed by each glucose measure (Sacks DA. Diabetes Care March, 2012). We have now discussed more the possible causes of heterogeneity in the current 3d paragraph in the Discussion.

Although lack of blinding to glucose values in most studies could be an explanation for the heterogeneity, we feel it less likely for outcomes other than Caesarean delivery, since we excluded treated women. In the EBDG study, clinicians did not receive routinely the results of the 75g OGTT, as they followed their usual clinical practice. The results of the 75g OGTT were part of the research protocol and were not routinely distributed with the argument that validated thresholds for this particular test were not available. Physicians could request the results and we had no control over this. We had the charts to abstract the interventions implemented. We have now discussed this in the current 3d paragraph in the Discussion.

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.

Response
Although this is an interesting suggestion, it departs from the main purpose of this paper, which was to provide a broad evaluation of both criteria in different settings using all available data.
Looking at Figures 7 and 8, we can make some initial comparisons. Results obtained for the EBDG showed always larger associations for the WHO than for the IASDP SG criteria, which are logically expected, since the IASDPG identifies lesser degrees of hyperglycemia. However, smaller associations for the WHO criteria were always seen in the HAPO study. We prefer not to single out this individual comparison, as the same phenomenon is generally present in the aggregate data and we believe that the individual comparison does not add much to the extensive discussion of heterogeneity which we are now doing in a more focused way in the Discussion section (3d paragraph).

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

Response
Thank you for pointing these out. We have made the corrections in Abstract and Methods.