Comparative Effectiveness Research Review Disposition of Comments Report

Research Review Title: Screening and Diagnosing Gestational Diabetes Mellitus

Draft review available for public comment from December 16, 2011 to January 13, 2012

Research Review Citation: Hartling L, Dryden DM, Guthrie A, Muise M, Vandermeer B, Aktary W, Pasichnyk D, Seida J, Donovan L. Screening and Diagnosing Gestational Diabetes Mellitus. Evidence Report/Technology Assessment No. 210 (Prepared by the University of Alberta Evidence-based Practice Center under Contract No. 290-2007-10021-I) AHRQ Publication No. 12(13)-E021-EF. Rockville, MD: Agency for Healthcare Research and Quality. October 2012. Available at: www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Comments to Research Review

The Effective Health Care (EHC) Program encourages the public to participate in the development of its research projects. Each comparative effectiveness research review is posted to the EHC Program Web site in draft form for public comment for a 4-week period. Comments can be submitted via the EHC Program Web site, mail or email. At the conclusion of the public comment period, authors use the commentators' submissions and comments to revise the draft comparative effectiveness research review.

Comments on draft reviews and the authors' responses to the comments are posted for public viewing on the EHC Program Web site approximately 3 months after the final research review is published. Comments are not edited for spelling, grammar, or other content errors. Each comment is listed with the name and affiliation of the commentator, if this information is provided. Commentators are not required to provide their names or affiliations in order to submit suggestions or comments.

The tables below include the responses by the authors of the review to each comment that was submitted for this draft review. The responses to comments in this disposition report are those of the authors, who are responsible for its contents, and do not necessarily represent the views of the Agency for Healthcare Research and Quality.

Source: http://effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1295
Published Online: November 5, 2012

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	General	This is a well written, comprehensive review of a difficult and controversial topic that has significant clinical implications. The authors have done an admirable job by assimilating the current literature and clearly stating 5 important topics related to the field. The key questions are clear and unambiguous.	Thank you for your comment.
Peer Reviewer #1	General	Some women cannot tolerate any form of GTT or OGS due to intolerance to oral glucose solutions. It would be nice to see this topic addressed somewhere in the paper. In the future, there is a need to find a screening or diagnostic methodology without side effects that could be utilized in general populations.	We have added a statement to the introduction of the main report.
Peer Reviewer #1	General – Clarity & Usability	Yes to all. I think the paper is exceptionally clear and organized (compared to many that I have studied or reviewed). The conclusions will be useful to those making policy or practice decisions but the delineation of knowledge deficits in the paper are particularly important for all to note.	Thank you for your comment.
Peer Reviewer #2	General	This report is a comprehensive review of the literature pertaining to the diagnosis of gestational diabetes mellitus (GDM). The authors aimed to identify test properties for GDM screening and diagnostic tests, evaluate potential benefits and harms of conventional screening (24 weeks and beyond) and early testing (<24 weeks), and evaluate the treatment of GDM with regard to changing outcomes.	Thank you for your comment.
Peer Reviewer #2	General	The authors at the Evidence-based Practice Center along with their consultants and technical expert panel identified 5 key strategic questions to meet their aims. The amount of work was massive and serves to highlight the difficulties in evaluating screening and testing for an entity for which there is no gold standard for diagnosis. While the report in and of itself will not solve the problem regarding the optimum screening and testing strategy, it provides a rich context to inform the debate.	Thank you for your comment.
Peer Reviewer #2	General	I have never performed a review of any manuscript that has left me without comments or criticisms. The authors are to be congratulated for admirably taking on this difficult subject.	Thank you for your comment.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #2	General – Clarity & Usability	The authors did an admirable job of organizing this body of evidence given the complex combinations of screening and diagnostic tests. The topic is inherently confusing than the authors do their best to synthesize the literature into a digestible form. In the end, as stated above, the report will provide valuable context for what will surely be a difficult discussion regarding diagnostic criteria for this enigmatic condition.	Thank you for your comment.
Peer Reviewer #3	General	While the report is technically competent and the key questions appropriately defined, the discussion lacks focus for the target audience of clinicians and health care providers. There is considerable repetition of results in the discussion without much analysis and the executive summary is far too long. For example, the problem of the lack of a "gold standard" for diagnosis of GDM is mentioned but no attempt to suggest what this might be and how studies might be designed to address this (although there are many excellent suggestions for research).	We have tried to reduce the repetitiveness of the discussion. We have kept the Executive Summary similar in length as this complies with AHRQ guidelines and could currently be considered a stand-alone document for a reader who does not have the time to read the full report. We have presented the evidence but we feel that it is the role of the stakeholder group to agree on a gold standard. This needs consideration of trade-offs between levels of sensitivity and specificity. Some methodologically rigorous studies have attempted to define a gold standard (i.e., specific threshold) and have been unable to do so; we have reviewed these studies and mentioned them specifically in the discussion.
Peer Reviewer #3	General – Clarity & Usability	As discussed above I believe the discussion is underdeveloped to inform policy & practice decisions. The Executive summary needs to be shortened and the extensive repetition between this, the main report and the discussion reduced. The main points are well presented.	We have made revisions to the discussion to highlight the evidence that might be used to inform decisionmaking. As mentioned above, it is beyond our remit to make recommendations, and the intent is that this report provides the evidence to inform discussions by the stakeholder groups. The Executive Summary is within the limits set by the publication guidelines of AHRQ.
Peer Reviewer #4	General	This review is comprehensive and does a good job summarizing the varied screens and blood glucose thresholds and the varied populations in which this has been studied. The 5 questions are clearly defined and the studies used for each question is clearly defined.	Thank you for your comment.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #4	General – Clarity & Usability	Well organized and well structured except for the issue of maternal weight gain. The report mentions that in some studies treatment of GDM benefited maternal weight gain but it is not clear if this is less or more maternal weight gain. For the obese, curtailing excessive weight gain would be an advantage while inadequate weight gain could be considered detrimental and improvement would be considered more weight gain. This needs to be more clearly reflected in the document.	We have clarified that maternal weight gain is a negative outcome (given the mean BMI of the study population). For additional information on this topic we refer you to a recent AHRQ report on Outcomes of Maternal Weight Gain (URL: http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?productid=528&pageaction=displayproduct
Peer Reviewer #5	General	This report is clinically meaningful addressing important clinical questions relating to the very common worldwide practice of screening for GDM. The fact that this report was commissioned by the AHRQ and intended to "assist and guide individual health plans, providers, purchasers and the health care system as a whole". It is anticipated that OMAR will use this data to inform a consensus meeting and develop guidelines. The key questions are clearly defined throughout the report.	Thank you for your comment.
Peer Reviewer #5	General – Clarity & Usability	The report is very well written and present the evidence in a clear and well organized manner. I would comment that there appears to be significant duplication in content between the very extensive executive summary and the body of the report itself. The conclusions, due to the limitations of the available evidence, will be difficult to use to create evidence based guidelines on screening for GDM. Again it seems that screening strategies and diagnostic thresholds for GDM will again be determined by consensus.	Thank you for your comment.
Peer Reviewer #6	General	This review was a massive task of pulling together a large and diverse literature. Overall you have done a thorough and masterful job.	Thank you for your comment.
Peer Reviewer #6	General	Given the large number of studies, it would help the reader to have a few summary sentences on your findings at the end of each KQ section both in the Executive summary and the main body of the report. You do have a bit of a summary at the very end, but there is a lot to wade through to get there. For a reader who is particularly interested in a specific KQ having this summary for each one would be helpful. This would also provide some clarity at the end of each section before moving the next for those of us wanting to plough through the entire work.	No change. We present summary Key Points near the beginning of each key question in the main report. We present a summary table of results in the Executive Summary and at the very end of the document.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #7	General	I have just been able to get online briefly. I read the exec summary and thought it was superb. I have no specific criticisms or suggestions.	Thank you for your comment.
Peer Reviewer #8	General	The EHC report entitled Screening and Diagnosis of GDM is clinically important as it searches the literature, assesses the quality of the literature and synthesizes the extensive data in the literature to provide evidence for address 5 key questions and fill in gaps since the last USPSTF review and demonstrate gaps that remain. This is particularly important as recent well designed studies have become the basis of institutional changes in the field of diagnosis and treatment of gestational diabetes. Yet, there remains considerable debate between institutions (e.g. American Diabetes Association versus American College of Obstetrics and Gynecology leaving most health care providers confused. The target audience is all interested health care providers in this field, as this is a public document and the Office of Medical Applications of Research for a consensus meeting and guideline development. Five key questions are clearly stated.	Thank you for your comment.
Peer Reviewer #8	General – Clarity & Usability	This was an outstanding study and with a high level of detail provided in the report and appendices. The conclusions can be made to inform policy and/or practice decisions. It is my opinion that recommendations provided in this report and in the future NIH consensus statement should strongly take into account the concerning increasing prevalence of childhood obesity. It is also my opinion that consideration should be given to using a method for diagnosing GDM that is linked to strong pregnancy and neonatal outcomes data.	Thank you for your comment.
Peer Reviewer #9	General	The report is very clear, logically structured and technical. The methodology and analysis is likewise clear, consistent in approach. It is a thorough review of existing studies, using a high standard for scoring the evidence.	Thank you for your comment.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #9	General	Sadly, the report does not provide "real world" direction to much needed answers. Using very high criteria for Quality Assessment of Studies (lots of low and insufficient in face of significant RR) should have more commentary (or this is for clinical discussion).	AHRQ EPC reports are meant to present the evidence, and not to make clinical recommendations. Making recommendations involves weighing benefits and harms and considering other individual values and resources. We hope that the EPC report will provide the evidence base to our partners to consider when making their individual or policy decisions or recommendations. The Limitations section discusses some of the reasons for low and insufficient strength of evidence assessments. We have offered some suggestions to address these limitations in the future research section.
Peer Review #9	General	In view of the lack of "gold standard" for diagnosis of GDM, I believe there should be a recommendation to utilize the 75 gm OGTT in all future studies to allow comparisons and base intervention studies on one common test (start with HAPO as a well designed foundation). Otherwise if everyone follows the recommendations for further "high quality" trials--it will not be possible to compare outcomes. We will be further down the road of confusion with GDM.	AHRQ EPC reports are meant to present the evidence, and not to make clinical recommendations. We have incorporated this suggestion into a recommendation for future research.
Peer Review #9	General – Clarity & Usability	The report is well structured and organized. This report clearly shows the lack of uniformity in diagnosis GDM and shows the continuous relationship between glucose and outcomes. Therefore it is critical for this report to recommend that ONE TEST be used internationally. If countries, areas decide different cut points for diagnosis but the same test--data can still be extracted similar to the comparisons between CC, NDDC, false positive screen, etc. The HAPO data provides a solid, high-quality basis to build quality future studies upon. Follow-up of the offspring will become available for some of the sites. We need to move forward, decide on one test (none will be superior because the relationship is continuous).	AHRQ EPC reports are meant to present the evidence, and not to make clinical recommendations. Making recommendations involves weighing benefits and harms and considering other individual values and resources. We hope that the EPC report will provide the evidence base to our partners to consider when making their individual or policy decisions or recommendations.
Peer Reviewer #10	General	The key questions pre-determined by OMAR and the USPSTF are relevant and more representative of the interrogations of women and caregivers than those of the previous 2008 review.	Thank you for your comment.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #10	General	The analytic framework is comprehensive but could be more clear and complete if one separates maternal and fetal short term (perinatal and pregnancy outcomes) from long term maternal/fetal metabolic effects.	The analytic framework was developed in consultation with the technical expert panel, OMAR, and AHRQ. No change.
Peer Reviewer #10	General	There are distinct direct and indirect links between screening with or without treatment and short vs long term outcomes. Long-term risks are not directly related to short term outcomes. Current screening strategies may not be cost-effective unless prevention of future maternal diabetes is achieved (Werner model, D Care; vol 35 march 2012). While this may look trivial, it is important because it emphasizes the difference between proponent of aggressive screening/ diagnosis/treatment of GDM (to prevent long term obesity/ diabetes) and those who considers the short term obstetrical impact.	We agree that there should be a distinction between short-term and long-term outcomes. We did not find evidence regarding future maternal diabetes and cannot comment on this outcome. We have reported on long-term outcomes among the offspring although the evidence was very limited. We have mentioned these results in our discussion and the need for more research and long-term follow-up.
Peer Reviewer #10	General	The authors should bring the issue of changing the definition of 'mild GDM' as a 'risk factor' of adverse outcomes instead of as the classic definition of a 'disease' characterized by glucose intolerance with onset or first recognition during pregnancy. It should be emphasized that the dichotomic view of GDM as present or absent is no longer acceptable because of the continuous relationship between maternal BG and outcomes so that the 'threshold issue' and the way to define it is a false problem. Maternal BG is for pregnancy as cholesterol is for CVD and perhaps different thresholds should be defined for different women risk.	We have added this comment to the discussion and conclusion sections.
Peer Reviewer #10	General	The absence of references in many parts of the manuscript (including tables and appendices) makes it difficult for the reviewer to verify the information.	We have reviewed the document and have added missing references to tables and the body of the report.
Peer Reviewer #10	General	Observational studies are often not easily distinctable from RCTs ; the latter should be given more emphasis	We have indicated the study designs (RCT, cohort) in the text and tables of the report.
Peer Reviewer #10	General – Clarity & Usability	The report is well structured but very long and at times repetitive.	Thank you for your comment. We agree that the report is long and repetitive but this is consistent with the format and technical aspect of these evidence reports. The Executive Summary provides a synthesis for a reader who is not interested or does not have the time to read the full technical report.
Peer Reviewer #3	Executive Summary	Executive summary is far too long (as stated in general comments).	No change.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #5	Executive Summary	The executive summary functions well as a stand alone document and will be what I suspect most people will turn to.	Thank you for your comment. We agree that many people will rely on the Executive Summary; therefore, we have not reduced the length as suggested above.
Peer Reviewer #5	Executive Summary	ES-1 line 34: should be East and South Asians instead of the blanket term of "Asians".	No change. The paper that we reference for this statement uses “Asians”.
Peer Reviewer #5	Executive Summary	ES-1 line 47: instead of "opposite effect of insulin..." would suggest using "hormones with anti-insulinic properties"	No change. We think our version is clearer.
Peer Reviewer #5	Executive Summary	ES-3 line 16 : the fasting plasma glucose level is not mentioned.	This information is available in Table 1, which is now included in the ES.
Peer Reviewer #10	Executive Summary	The choice of specific 'key outcomes' (different from 2008 report) and especially the post-hoc decision to include shoulder dystocia and macrosomia should be discussed further by authors (page ES-8, lines 41 to 43)	The decision to include shoulder dystocia and macrosomia was not post-hoc. These were included in the key questions. The post-hoc decision we made was to grade these outcomes. The clinical leads felt that is was important to provide strength of evidence assessments for these important clinical outcomes. We have commented in the methods section on why we decided to grade these outcomes post-hoc.
Peer Reviewer #10	Executive Summary	Rregarding definition of GDM not included	No change. We are not sure what this refers to.
Peer Reviewer #10	Executive Summary	Page ES-1, lines 27:'...IADPSG ...in which lower glucose thresholds are accepted' should be change by... are proposed	We have changed the wording in the ES and main report to read “…in which lower glucose thresholds are proposed to diagnose GDM.”
Peer Reviewer #10	Executive Summary	Page ES-1, lines 21 to 23: The affirmation stating that '...data from HAPO indicates that 7% of women met an FPG of 5.3 and over...' seems erroneous. There was 2.9% women with FPG 5.3 to 5.5 and 0.9% with FPG 5.5. to 5.8 mmol/l (total 3.8% (plus max 2.9% who were excluded because of higher BG, either fasting or 2h	We have confirmed that our number is correct for women (blinded or unblinded) with a FPG threshold of ≥5.3 mmol/L.
Peer Reviewer #10	Executive Summary	Page ES-1, lines 32-33: authors cite ref 6 but should mention that these prevalence of GDM in the US was estimated by IADPSG criteria (line 30 'prevelence' is wrongly spelled	We have made this change in the ES and main report. We have corrected the spelling of prevalence in the ES and main report.
Peer Reviewer #10	Executive Summary	Page ES-1, lines 36 to 39: authors cite ref 7 to mention ethnic difference in GDM prevalence; for clarity, they should add that GDM prevalence was estimated according to Carpenter and Coustan criteria and/or hospital discharge diagnosis	We have incorporated this change to the ES and main report.
Peer Reviewer #10	Executive Summary	Page ES-2; Screening and diagnostic strategies should refer to Table 1 (Diagnostic criteria and plasma glucose thresholds for gestational diabetes mellitus on page 5	We have included a copy of Table 1 in the ES and have referred readers to it.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #10	Executive Summary	Page ES-2 line 42: the ref 13 is wrongly printed	Reference 13 in the ES is correct. No change.
Peer Reviewer #10	Executive Summary	Page ES-3, lines 18-19: ...diagnosis of GDM is made when one or more values (should be 2 or more BG values	This comment is incorrect. IADPSG only requires 1 abnormal glucose.
Peer Reviewer #10	Executive Summary	Page ES-3, lines 50:'...that identified a 1.75-fold increase in large for gestational age...' should add '...adjusted odd ratios (aORs) of 1.75 relative to the mean cohort glucose values...' 1.75-fold is a relative risk (RR) that is not always equivalent to an OR of 1.75 (OR overestimates the RR if prevalence of outcome studied is 10% or over)	We have incorporated this change in the ES and main report.
Peer Reviewer #10	Executive Summary	Page ES-11 line 45: prevalence of GDM according to ADA (75g) 11.2 to 19%...' ; authors should clarify which ADA criteria were used.	We have included this information in the ES and main report: ADA 2000-2010.
Peer Reviewer #10	Executive Summary	Page ES-15, lines 37 to 39: authors consider no difference in maternal weight gain because they exclude Crowther RCT that showed less weight gain in treated women (from the first prenatal visit instead of from enrolment). They should explain this choice more clearly in pages 65-66 and figure 48	Thank you for pointing this out. We have included the data from the Crowther study. We did not pool the results because of substantial heterogeneity. Two of the RCTs showed no difference, while two RCTs (including Crowther) showed a significant difference. We have added this detail to the text of the results section.
Peer Reviewer #1	Introduction	Nice summary of the state of the art and existing controversies. Particular attention is given to different diagnostic criteria and population differences, topics of great importance to the clinician.	Thank you for your comment.
Peer Reviewer #2	Introduction	The Introduction provides an excellent review of the epidemiology and clinical issues associated with GDM, and leads the reader in a logical fashion to the rationale for the key questions to be addressed by the report.	Thank you for your comment.
Peer Reviewer #3	Introduction	There is considerable duplication between the executive summary introduction & the main document introduction that i found irritating but may be part of a "house style".	No change. We agree there is duplication between the ES and the main report; however, the ES is written to be a standalone document.
Peer Reviewer #4	Introduction	Appropriate.	Thank you for your comment.
Peer Reviewer #5	Introduction	The authors have provided a thorough and up to date review of the current state of the literature as well as identifying the significant areas of controversy regarding GDM screening.	Thank you for your comment.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #8	Introduction	The introduction is clear and summarizes background, risk factors, screening and diagnostic strategies and problems related to lack of a "gold standard", as well as important new studies. The introduction also notes treatment strategies, specifically diet, insulin and oral antidiabetic medications (ES 4). The strategy for this review is described and the key questions are clearly defined. I would like to point out that while the introduction notes the use of oral agents for the treatment of GDM, this is not discussed in the remainder of the report and there is no data comparing oral agent versus insulin therapy. The use of oral agents in GDM is controversial and a review of this literature would be very helpful.	The use of oral agents in the treatment of GDM is beyond the scope of this report. We agree that this is an important question and refer readers to the systematic review by Waugh et al (Health Technology Assessment 10:14 (45).
Peer Reviewer #9	Introduction	A clear and concise summary of the various diagnostic criteria for GDM. The questions are systematic in structure and have appropriate and meaningful outcomes.	Thank you for your comment.
Peer Reviewer #1	Methods	Inclusion and exclusion criteria are reasonable. In fact, I applaud the attempt to exclude suspected pregestational diabetics from the analysis. Definitions are clear. Diagnostic criteria for outcomes seems appropriate. Potential limitation and weakness are acknowledged; for example, the authors note that foreknowledge of the existence of GDM might influence decision to perform c-section.	Thank you for your comment.
Peer Reviewer #2	Methods	The literature search was exhaustive and appropriate. Although some might argue that literature prior to 1995 might be included, this reviewer very much agrees with including this relatively contemporary and yet large body of literature. Studies were included based on their ability to shed light on the key questions. Study quality was assessed and strength of evidence for questions 4 and 5 (questions addressing the effect of treatment and potential harms of treatment) was considered. Statistical methods were straight-forward and appropriate for addressing the key questions.	Thank you for your comment.
Peer Reviewer #3	Methods	I believe the inclusion & exclusion criteria are appropriate, the search strategies, statistical measures & outcome measures are appropriate.	Thank you for your comment.
Peer Reviewer #4	Methods	The identification and scoring of each source is well outlined.	Thank you for your comment.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #5	Methods	The methods are clearly stated and appropriate for this type of report. The decision to limit the review to randomized controlled trials and cohort studies is stated but the reason for not including case control studies (although understandable) is not discussed and this should be addressed. The outcome measures are listed but not individually defined in the methods. The outcomes are defined in the results section when analyzing individual study results.	No change. Given that there are RCTs and cohort studies that address the key questions, the research team (in consultation with the technical expert panel and AHRQ) determined that the inclusion of case-control studies was unnecessary. This was specified in the study protocol developed at the outset of the project.
Peer Reviewer #8	Methods	The literature search was explicitly stated and comprehensive. The inclusion and exclusion criteria are justifiable. Definitions and diagnostic criteria are appropriate. Statistical methods seem appropriate, though I would defer to a statistician on this, as the methods seemed quite advanced and beyond my level of understanding.	Thank you for your comment.
Peer Reviewer #9	Methods	The inclusion/exclusion were justifiable and the search strategies were exhaustive (well state and in depth). The outcome criteria were clinically useful and appropriate. The statistic methods were appropriate. As I am not a statistician the Quality Assessment seemed impressive and correct. However I am concerned that the reviewers deal in a perfect world not the "real world".	Thank you for your comment.
Peer Reviewer #10	Methods	A statement saying why a random effect model was chosen for the meta-analysis would be appreciated	We have added a statement to the methods section of the main report.
Peer Reviewer #10	Methods	The meta-analysis of RCTs for treatment is dominated by 2 studies; this fact should be discussed.	We have added this point to the key points within the results section of Key Question 4, as well as to the discussion section of the ES and main report.
Peer Reviewer #10	Methods	The Landon RCT 2009 study (MFMU) has a negative primary outcome. It should be mentioned that the secondary outcomes have to be regarded as exploratory findings.	When conducting a meta-analysis we do not take into consideration whether an outcome was a primary or secondary in the relevant study. We enter the data for each outcome as presented in the primary study and pool the results from all relevant included studies. What is more important in the context of a meta-analysis and systematic review is the priority of outcomes specified by the review team.
Peer Reviewer #1	Results	Overall, no issues here.	Thank you for your comment.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #2	Results	The Results section of the report is laid out logically addressing each key question. Included studies are described briefly in the narrative and in more detail in the appendices. Results are succinctly summarized and displayed with appropriate graphics.	Thank you for your comment.
Peer Reviewer #3	Results	I am not aware of any studies that have been missed & they are tabulated and described appropriately. I particularly like the tables that show the effect of prevalence e.g. page 23ES. The detail provided is clear and sufficient.	Thank you for your comment.
Peer Reviewer #4	Results	Appropriate.	Thank you for your comment.
Peer Reviewer #5	Results	The details presented in the results are extensive and appropriate. The use of "key points" makes the data more manageable for the reader.	Thank you for your comment.
Peer Reviewer #5	Results – KQ1	Regarding the tables and figures - I am not sure the HSROC curves (KQ-1) add to the data presented in the tables and Forest plots. In general it would be useful to add the reference numbers to the studies quoted in the tables and forest plots. In the chapter addressing risk-factor based screening (page 32 line 27) it is not clear to me why the following study was omitted: Griffin ME, Coffey M, Johnson H, Scanlon P, Foley M, Stronge J, O'Meara NM, Firth RG. Universal vs. risk factor-based screening for gestational diabetes mellitus: detection rates, gestation at diagnosis and outcome. Diabet Med 2000 Jan;17(1):26-32 This is a rare RCT enrolling women prior to screening for GDM and provides data regarding key question 1. It is not listed as an excluded study. Clarification is required.	We have moved the HSROC curves to the appendix. The Griffin study did not meet our inclusion criteria as it did not provide data for a 2x2 table. It is listed in our table of excluded studies.
Peer Reviewer #5	Results – KQ2	Regarding key question 2 I am not sure that the 2 studies included correctly address this key question. The authors correctly state in the discussion that to answer this question would need to be identified in a RCT randomizing women to screening vs no screening. The authors also correctly address the limitations of the two included studies which are retrospective and unable to control for key confounders. Although I was unable to access the study from Thailand (Ref 120) I think that it would be more accurate to simply state that there were no studies identified that address this key question.	These 2 studies met our a priori inclusion criteria. We feel that our summary statements about the impact of screening being inconclusive accurately reflect the state of the evidence.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #5	Results – KQ3	Regarding the results for key question 3 there is a significant limitation due to the heterogeneity of the studies with regards to diagnostic criteria and the comparison group used (GDM according to different criteria, non GDM, GCT false positives etc.). Obviously comparing women with varying degrees of glucose intolerance to the "normal" population will yield different results than when comparing outcomes to an at risk population. Adjustment for treatment and clinician bias when knowing the screening/diagnostic results also limits the ability of the majority of the studies included to answer this question accurately. This should be further discussed in the limitations.	We have incorporated this point into our limitation section in the ES and the main report.
Peer Reviewer #5	Results – KQ4,5	For key question 4 the RCTs used are not of equal quality and differ significantly in the population randomized. In reality only the Crowther and Landon RCTs should be included as they provide the best evidence to answer this criteria. Of the three additional RCTs two of the RCTs randomize women with false positive GCTs and thus don't address key question 4 specifically.	All the studies in this section met our inclusion criteria. As appropriate, we have highlighted the findings of these 2 trials.
Peer Reviewer #6	Results	Considering the findings not just as a list but within the clinical context would be very helpful. Are there ways the clinician might discern a path through this dense forest? Without making recommendations, you can make clinically relevant observations and organize findings that will help those formulating guidelines and others to fully comprehend the evidence.	We have revised the discussion and attempted to highlight the evidence that may be helpful to decisionmakers.
Peer Reviewer #6	Results	Are there ways to group (not pool, but consider findings in a qualitative way) studies getting at, for example, outcomes of treatment of GDM or of untreated GDM. Just saying that a single study always means the evidence is insufficient is classically correct, but when the differences in the studies were primarily in the diagnostic criterion used to identify GDM and outcomes were similar or related, might there still be something to learn from the pattern of the evidence?	We have added the following statement to our discussion of the KQ3 results which we feel addressed the reviewer's comment: "While many studies have attempted to measure the association between various criteria for GDM and pregnancy outcomes in the absence of treatment, the ability of a study or pooled analysis to find a statistically significant difference in pregnancy outcomes appears more dependent on study design, in particular the size of the study or pooled analysis, rather than the criteria used for diagnosing GDM."

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #6	Results	In general, when can exploration of the possible underlying clinical (not just statistical) contributors to heterogeneity tell us? At several points you comment about a single study being responsible for statistical heterogeneity, but are there any discernable differences in the population or study protocol that will shed light on the differences?	We have corrected an error identified by another reviewer which has reduced the statistical heterogeneity that was being caused by the single study. Where statistical heterogeneity is high, we explored potential explanatory clinical variables but no clear patterns emerged. We have added text to the results explaining this.
Peer Reviewer #6	Results – KQ1	In KQ1 there are many comparisons, using many standards. Given this diversity of evidence, are there any useful threads, or do we just say there is so much diversity we can't really conclude much based on 44 eligible studies. Based on your findings, what test or tests will be most likely to optimize sensitivity and specificity (a common goal). Would this vary based on the population being tested? i.e., high or low prevalence, other factors? How do these screening tests relate to current practice in the US and Canada? How do they relate to the various guidelines? At my institution, all pregnant women are screened in the first trimester and get the 2 hour 75 gram glucose test at 24-28 weeks, according to the the IADSG criteria. As best I can tell, there is one study to support the 75 gram 2 hour test as a single test, and really not much evidence at all to support screening for DM (or how it should be done) in the first trimester. Is this correct? If so, statements linking existing evidence to the various guidelines would be helpful for those of us who feel that we are swimming in a sea of alphabet soup.	The role of the EPC and this report is to present the evidence regarding test characteristics of screening criteria. Stakeholder groups will need to assess the balance between sensitivity and specificity to make recommendations regarding specific tests. We have attempted to clarify the relationship between the different criteria – see Table 1 and Figure 1. Bibliographic references are provided.
Peer Reviewer #6	Results – KQ3	In KQ 3, I found the graphics a bit confusing for 2 reasons: a) I would prefer to see a label specifying the outcome under consideration on the x axis. This would help to clarify what the RR relates to. "Favors false positive" vs "Favors No GDM" or "Favors GDM" vs "Favors No GDM" are confusing labels when there is a potentially elevated RR for mortality, for example. I puzzled over these for several minutes before I more or less sorted out what the graphic was getting at. Using these labels as subheadings under a clear label about the outcome of interest would be more helpful.	The format of the forest plots is the standard way to present the data. We have stated the outcome in the title of each figure. We have tried to provide interpretation of the results in the text that precedes each forest plot.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #6	Results – KQ4	Does it really make sense to use a summary RR symbol just below the box and whiskers when there is only a single study in the subgroup under consideration? (see for example pp. 69 and 70 of the main report). Having 2 symbols there relating to one study is a bit confusing and potentially misleading.	We realize this can be confusing, however, this is the standard output for the meta-graphs. In all meta-graphs where there is only 1 study for a subgroup, there is more than 1 study for the other subgroup. We cannot remove the summary symbol without removing it from both subgroups.
Peer Reviewer #8	Results	There is very extensive, helpful detail in the body of the report and in the appendices.	Thank you for your comment.
Peer Reviewer #8	Results – KQ3	Maternal outcomes/preeclampsia (page 43)- consider including data from Metzger et al IADPSG Recommendations on the Diagnosis and Classification of Hyperglycemia in Pregnancy Diabetes Care 33: 676-672 and especially the on-line appendix Table B which compares frequency of preeclampsis when all blood glucoses are below the thresholds versus when 1 or more are at or above the thresholds. It would seem that this data might strengthen the evidence for an association between increasing glycemia and preeclampsia.	Thank you for this information. We have incorporated the data into our results, as appropriate.
Peer Reviewer #8	Results – KQ3	Fetal/Neonatal Outcomes- macrosomia the study by Hillier et al ref 122 was eliminated because it was highly "influential" and resulted to heterogeneity of the data. Yet the Hillier study was rated 9/9 by the Newcastle-Ottawa Quality Assessment Scale. I looked at the data from the paper and it appears that the number of events (macrosomia >4000 gms) was collected incorrectly. I believe that the data for <4000 gms was collected in place of >4000 gms. Therefore, for figure 25 events should be 25 instead of 148 for GDMs and 905 instead of 6695 for No GDM. Similar mistakes were made in figures 26 and 27.	Thank you for pointing this out. We have made the correction.
Peer Reviewer #9	Results	The detail is appropriate. The tables and graphs are very helpful. No studies were exclude to my knowledge.	Thank you for your comment.
Peer Reviewer #9	Results	Again--KQ3 has lots or significant RR despite the "low" quality.	We have addressed the reasons for low and insufficient SOE assessments in the limitations section.
Peer Reviewer #10	Results – KQ4	Page 80, line 36: ref 148 is not related to fasting glucose as a test but is a commentary article	We have removed this reference.
Peer Reviewer #10	Results – KQ4	Pages 86 line 54 and 87 lines 3 and 4, the same comment applies to avoid the term 'risk' and replace by adjusted odd ratios as it is not equivalent i.e.:'... 1.75 risk of LGA... or 1.4 risk for pregnancy hypertension...' (see point 8 of the introduction comments above	This change has been made in the ES and main report.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #10	Results – KQ5	Key question 5, as for benefit, I suggested to separate short and long term risks of harm in mother and infants and consider long term potential increase/decrease in costs and resources for healthcare if all GDM women are followed to prevent diabetes	We have specified which are short-term and long-term outcomes when we present the results for KQ5. The studies that met the inclusion criteria did not report on long-term resource use.
Peer Reviewer #10	Results – KQ5	Key question 5 should also address potential long term harm of being labeled as GDM on postpartum follow-up, decision to have another pregnancy and management of future pregnancies (many centers consider them as GDM but the 2 dominant RCTs of treatment excluded women with previous GDM), maternal insurability (no data but important to mention, perhaps in the future research section page 89)	We have incorporated this suggestion into the discussion and future research sections in the ES and main report.
Peer Reviewer #10	Results – KQ5	Page 87, lines 12-13: 'A change in diagnostic criteria without addressing management thresholds could contribute to clinical confusion.' This sentence is too moderate. Any change in diagnostic criteria will most probably lead to change in management/BG targets of women with GDM. Recommendations on new thresholds that do not address the issue of treatment are not only 'confusing' but irresponsible and potentially dangerous. If treatment change is to be made, it should be in a research setting in a RCT with ethical approval and women consent.	We have left this sentence as is, but incorporated the reviewer's concerns in the following sentence: "If diagnostic thresholds for GDM below the treatment targets of the large RCTs are endorsed, this could ethically obstruct the possibility of future RCTs to compare different treatment targets above such diagnostic thresholds."
Peer Reviewer #10	Results – KQ5	Page 87, lines 29-30: 'The ongoing obesity epidemic in the United States warrants careful consideration of a diagnostic approach for GDM that incorporates maternal BMI.' This is a very important suggestion as maternal BMI/weight gain are more/equally important for adverse outcomes and clinicians do not rely on isolated OGTT relative risks or odd ratios but on women individual multifactorial absolute risks. Any model that incorporates maternal BMI as well as other modifiable risk factors should be validated in a formal RCT before being largely diffused.	We have added this point to the discussion: "This would require the development and validation of a risk model that incorporates maternal BMI as well as other modifiable risk factors."
Peer Reviewer #1	Discussion	As noted above, a fair acknowledgement of study implications is duly noted. The section on future research needs is particularly important. As I read reread portions of the paper, I had many of the same concerns and "burning questions". The discussion nicely summarizes many of the knowledge deficits that should be addressed in future research.	Thank you for your comment.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #2	Discussion	The authors appropriately summarize their findings and recognize the difficulties in diagnosis when the condition has no gold standard and the effects of treatment will likely vary depending on the thresholds chosen for diagnosis. They aptly point out that the continuous relationship between glucose levels and maternal and neonatal outcomes makes the establishment of a diagnostic threshold problematic and that research is needed to tie thresholds to benefits and harms of treatment.	Thank you for your comment.
Peer Reviewer #3	Discussion	This is where I have a problem with this otherwise excellent review. The list of research is clear but much of the discussion is tedious repetition of earlier parts of the report without sufficient development conclusions to inform policy. While part of this is a result of the messiness of the area under consideration, without this more thoughtful discussion the field will not progress. The challenges to this e.g. the difficulty of research in pregnant women should be explored and potential solutions (e.g. regional/national registers) outlined. The issues of the lack of a gold standard is not addressed and is critical to advancement of any of the other programmes of research.	AHRQ EPC reports are meant to present the evidence, and not to make clinical recommendations. Making recommendations involves weighing benefits and harms and considering other individual values and resources. We hope that the EPC report will provide the evidence base to our partners to consider when making their individual or policy decisions or recommendations.
Peer Reviewer #4	Discussion	Excellent covering limitations and need for more research.	Thank you for your comment.
Peer Reviewer #5	Discussion	Overall the discussion successfully summarizes the extensive results and presents the conclusions clearly and accurately. The limitations are clearly stated. The conclusions reflect these limitations. Future research avenues are clearly identified but I would add the following: 1) Further research should be directed at identifying, through a combination of clinical and biochemical risk factor, those at highest risk of adverse pregnancy outcomes associated with hyperglycemia in pregnancy. This would allow the health care system to focus efforts on a select group of high risk women without exposing lower risk women to unnecessary interventions. 2) Further investigation is required to separate the independent contributions of maternal adiposity versus glucose intolerance to adverse perinatal outcomes.	We have incorporated these points into our discussion; however, we have not included them as specific research needs because they don't stem directly from our key questions and results.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #8	Discussion	The major findings are clearly stated. In particularly: "evidence supports benefits of treating mild GDM with no evidence of harm. Specifically the treatment of GDM results in lower incidence of preeclampsia, macrosomia and LGA." I would again point out that risks and benefits and comparison of oral agents with insulin was not addressed despite its mention in the introduction. I agree with the importance of further study of the long term metabolic effects on the infants of diabetic mothers. I agree with the authors additional identification knowledge gaps.	No change. The issue of oral agents vs. insulin in the treatment of GDM was not part of the key questions addressed by this report.
Peer Reviewer #9	Discussion	The Key findings for the questions were information packed, covered everything with densely packed data. The summary tables are helpful (table 17 & 18). I would include one for KQ3 as well.	We have added a summary table to the discussion in the main report.
Peer Reviewer #9	Discussion	The studies referenced in "findings in relationship to what is known" sites the MOST USEFUL studies. Both deal with mild GDM (largely)--a very important point. These studies use different criteria but have largely the same outcomes. HAPO is a blinded prospective studies for the purpose of finding common ground. Most interesting to me is the larger effect of BMI on outcomes, including cord C Peptide. I strongly urge that this report recommend the adoption of one test (75g) (even if more analysis to decide cutpoints are undertaken). We need well done intervention and follow-up studies--but we must start with the same test at least.	We have incorporated this suggestion into our discussion and the future research needs.
Peer Reviewer #10	Discussion	It would be useful to discuss differences in methodology and findings between this review and the previous meta-analysis of RCTs (BMJ 2010;340).	Thank you for this suggestion. We have incorporated it into the discussion section.
Peer Reviewer #10	Conclusion	Conclusion page 90, line 15 and lines 18-21: I suggested to precise '...no evidence of short term harm	We have made the change.
Peer Reviewer #10	Conclusion	Conclusion page 90, line 17: '...large for gestational age.' add the term infants or neonates	We have made the suggested edit.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #10	Conclusion	Conclusion page 90, lines 22 to 26 should be less moderate: 'What remains less clear is which diagnostic thresholds for GDM should be chosen. Given the continuous association between glucose and a variety of outcomes, decisions should be made in light of what outcomes that are altered by treatment are most important and what level of increased risk is acceptable.' The only rigorous conclusion should be that, given the continuous association between BG and outcomes, diagnostic criteria proven to modify outcomes in RCT treatment studies should be used. Because the primary outcome in the 2009 Landon RCT study was not attained, the only clear evidence to date are the diagnostic criteria used in the Crowther RCT.	Our report provides a summary of all of the evidence we identified for clinically important outcomes. We have presented the totality of evidence including the findings from the two large RCTs.
Peer Reviewer #10	Conclusion	Conclusions are generally fair	Thank you for your comment.
Peer Reviewer #10	Appendix	Table D-3, page D-22: line 12 of the Chen study, the n is probably 1469 instead of 14,69	Correction has been made.
Peer Reviewer #10	Appendix	Table D-4 page D-34, Crowther study: lines 17-18 the number under 'interventions' should be 5,5 and 7,0 mmol/l instead of 55 and 70 and it should be add that a BG target of under 8.0 mmol/l was set at more than 35 weeks of pregnancy; same thing line 31, the number under 'screening' should be 7,8 instead of 78 mmol/l	We have made the changes.
Peer Reviewer #10	Appendix	Table D-4 page D-37, Landon study: line14: ih cut off is missing (between 135 and 200 mg/dl or 7.5 and 11.1 mmol/l)	We have made the change.
Peer Reviewer #10	Appendix	The terms IMC, IWG and DPSG not always explained on the table D-1 legends	We have made the changes to the legends.
Peer Reviewer #10	Appendix	References page E-15: ref 25 and 26 are the same (Yogev 2003)	We have corrected the references.
Peer Reviewer #10	Appendix	pages F-3 to F-5: lines 19 and 31 refers to table E-1 and not table 1; same thing line 24, F-4 and line 15, F-5	We have made the corrections.
Betty C. Jung	General	Good comprehensive scientific documentation. This should provide evidence of the need for screening pregnant women to distinguish between gestational and type 2 diabetes.	Thank you for your comment.
Betty C. Jung	General	I would recommend that you provide a meta-analysis of whatever studies have been done on the use of H1Ac for the diagnosing of type 2 diabetes.	A review of diagnostic tests for type 2 diabetes was beyond the scope of the report.

Commentator & Affiliation	Section	Comment	Response
Betty C. Jung	General	When I was the state staff cardiovascular and diabetes epidemiologist for the Connecticut Department of Public Health, I chose gestational diabetes as a priority area for additional research and public awareness. To that end I wrote an issue brief on this (http://www.bettycjung.net/Inprint/GDM2008.pdf) Based on my research and continual interest about gestational diabetes, I believe that there is a major advantage to promoting the use of H1Ac in pregnant women to distinguish between those who eventually develop diabetes and those who already have type 2 diabetes, but were not diagnosed before they became pregnant. This can be easily accomplished by performing an H1Ac during the first 3 months of pregnancy. Since most women will not know for sure they are pregnant until they have missed at least 2 periods, this would mean they are about 2 months pregnant when they are sure with a positive pregnancy test. If an H1Ac is drawn at this time, then an abnormal reading would automatically identify the presence of type 2 diabetes (or pre-diabetes) in a pregnant woman, thus ruling out gestational diabetes in these women. In fact, this is in keeping with the American Diabetes Association's Standards of Medical Care in Diabetes – 2012	Thank you for your comment.
Betty C. Jung	General	Finally, I appreciate the AHRQ coming out with this draft document for public comment. I look forward to seeing this implemented so we can move forward with addressing the growing epidemic of type 2 diabetes. Addressing gestational diabetes and type 2 diabetes in pregnant women is keeping with Healthy People 2020 diabetes objectives and makes good common sense to intervene early to prevent the development of type 2 diabetes in those at greatest risk of developing the disease.	Thank you for your comment.

Commentator & Affiliation	Section	Comment	Response
American Diabetes Association	General	Despite the growing burden of GDM, in 2003 and again in 2008, the U.S. Preventive Services Task Force (USPSTF) concluded that the "evidence was insufficient to assess the benefits and harms of screening for gestational diabetes mellitus either before or after 24 weeks gestation,"[1] issuing an "I" recommendation regarding routine GDM screening for all pregnant women. Despite these recommendations, diabetes and obstetrical societies almost uniformly recommend screening all or most pregnant women for GDM, and studies have shown that over 95% of US obstetricians regularly screen for GDM.[2] Additionally, the Secretary of Health and Human Services recently adopted the Institute of Medicine's recommendation to include GDM screening in its list of preventive services for women that will be covered with no cost-sharing beginning in August 2012. We applaud the reviewers for considering the full body of available evidence—including randomized controlled trials; nonrandomized controlled trials; and prospective and retrospective cohort studies—in drawing conclusions about GDM screening.	Thank you for your comment.
Anonymous	Executive Summary	I am confused. Where are the data on the FBS + 1 hr & 2 hr postprandial with 75 gm load that we've been using for a couple of years now?	We are not sure what the reviewer is asking. Table 1 summarizes thresholds from various organizations. We have now included this table in both the main report and the ES. For Key Question 1, we have included studies that examined a 75 g load.

[1] www.uspreventiveservicestaskforce.org/uspstf08/gestdiab/gdrs.htm

[2] Gabbe S, Gregory R, Power M et al. Management of diabetes mellitus by obstetrician-gynecologists. Obstet Gynecol 2004; 103(6):1229-34.

Source: http://effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1295

Published Online: November 5, 2012

Commentator & Affiliation	Section	Comment	Response
Anonymous	Introduction	Gestational diabetes mellitus (GDM) is a form of diabetes developed during pregnancy. It affects over 10% of all pregnancies in the United States each year. The condition can lead to short- and long-term health problems for the mother and fetus including immediate problems during delivery and extended risks for developing insulin resistance and type 2 diabetes postpartum. This thesis is constructed as an accessible document within the public health and medical anthropology disciplines because of its focus on primary prevention, health education, and the biocultural realities that put people at risk for diabetes during pregnancy. This thesis aims to explore five major questions: (1) what is the significance of GDM within the broader diabetes epidemic in the United States; (2) what is the current approach to care surrounding GDM; (3) what are the shortcomings of current medical protocols; (4) what can be improved; and, (5) how can focusing on and introducing prevention to children born to a GDM environment provide a novel approach that intervenes in transgenerational transmission of risk factors. Type 2 diabetes is chronic disease that is developed, meaning it is a disease that can be prevented and managed via lifestyle modification including healthy nutrition, ample exercise, minimal stress, and not smoking. With GDM and type 2 diabetes rates dramatically rising in the United States, more people are living with the conditions or certain risk factors including poor diet, lack of exercise, family history and, therefore, pass the risks for diabetes along generational lines. Gestational diabetes mellitus and more generally the time during pregnancy offer important avenues to address the diabetes epidemic in all populations within the United States. The decisions made during pregnancy affect the immediate biological realities for mothers and newborns and can encourage new models of diet and activity that can promote a healthier lifestyle which could last a lifetime and may have important influence on redirecting health outcomes for future generations.	Thank you for these observations.

Commentator & Affiliation	Section	Comment	Response
Academy of Nutrition and Dietetics	Results – KQ1	The Academy agrees with the Committee on Obstetric Practice of the American College of Obstetricians and Gynecologists that "[a]ll pregnant woman should be screened for GDM, whether by patient history, clinical risk factors, or a 50-g, 1-hour loading test to determine blood glucose levels."2 In addition, all women diagnosed with GDM (after one or more plasma glucose values exceed established cutoffs) should receive nutrition counseling by a registered dietitian. The Academy notes the Summary and Recommendations of the Fifth International Workshop-Conference on Gestational Diabetes Mellitus that "MNT is best prescribed by a registered dietitian or qualified individual with experience in management of GDM."3 Screening based solely on known clinical risk factors for GDM, such as age, weight, ethnicity, and family history of diabetes fails to identify one-third to one-half of affected pregnant women. Recognizing the lack of international consensus regarding screening and diagnostic criteria, the Academy has encouraged the National Quality Forum to develop quality measures related to GDM screening and referral to dietitian for medical nutrition therapy.	Thank you for your comment.
Academy of Nutrition and Dietetics	Results – KQ2	Based upon available evidence, the Academy recommends that all pregnant women be assessed for risk of GDM at the first prenatal visit. Depending on level of risk, timing of screening for GDM and/or impaired glucose tolerance (IGT) during pregnancy will differ. The Academy recommends that an RD or other qualified member of the interdisciplinary team should advise women with GDM to monitor their blood glucose, including fasting and postprandial levels. Harms associated with screening women for GDM include a lack of international consensus of the screening and diagnostic criteria, the potential for false-positives, the potential of psychological stress for some individuals, and the possibility that screening may cause gastrointestinal upset and other symptoms in some individuals.	Thank you for your comment.

Commentator & Affiliation	Section	Comment	Response
Academy of Nutrition and Dietetics	Results – KQ3	The Academy notes unfavorable health outcomes of mothers who meet various criteria for GDM and their offspring in the absence of treatment. Research indicates the similarities between GDM and IGT during pregnancy, and both are associated with increased risks of poor maternal/neonatal outcomes if left untreated, including large for gestational age newborns, macrosomnia, increased risks of preterm birth, perinatal morbidity, and neonatal hypoglycemia.	Thank you for your comment.
Academy of Nutrition and Dietetics	Results – KQ4	Modification of nutritional intake is specifically mentioned in both Key Question 4 and Key Question 5 as an effective component of the treatment in the intervention groups. The Academy notes that although no specific type of nutrition intervention is highlighted, the two RCTs reported significantly more visits with an RD (92% vs 10%) than the control group.6 This is a significant finding, which strongly suggests that any type of intervention should include an RD. *Medical Nutrition Therapy Treatment* The available evidence leads the Academy to recommend RD-provided MNT initiated within one week after diagnosis of GDM, to include a minimum of three nutrition visits. Research, including the Reader study not included in the Evidence Synthesis,7 indicates that MNT is essential for demonstrably improved maternal and neonatal outcomes, especially when diagnosed and treated early.8 For women with IGT during pregnancy, the Academy recommends that RDs should initiate the same recommendations of MNT as those for women diagnosed with GDM. Research indicates that IGT and GDM carry similar risks of adverse outcomes. Medical nutrition therapy, with glucose monitoring, was the only therapy in >90% of women with gestational diabetes in 2 large RCT trials, which highlights the importance of referral to dietitian and for follow up visits to assess adherence and need for further therapy.	Thank you for your comments. The study by Reader et al (J Am Diet Assoc 2006) was excluded from the review because it did not have relevant comparisons. It is listed in our list of excluded studies in the Appendices.

Commentator & Affiliation	Section	Comment	Response
Academy of Nutrition and Dietetics	Results – KQ4	*Pharmacological Therapy* The Academy recommends that when optimal blood glucose levels have not been maintained with MNT and/or the rate of fetal growth is excessive, RDs should encourage the initiation of pharmacological therapy for treatment of women with GDM. Research indicates that although recommended target blood glucose levels vary among organizations, pharmacological therapy, such as the use of insulin, insulin analogs and glyburide, improves glycemic control and reduces the incidence of poor maternal and neonatal outcomes. Unless contraindicated, the Academy recommends that RDs should encourage breastfeeding in pregnant women, including those with GDM. Research indicates that even short duration of breastfeeding results in long-term improvements in glucose metabolism and may also reduce the risk of type 2 diabetes in children. Breastfeeding has been associated with decreasing the woman's risk of becoming overweight later in life and developing metabolic syndrome and type 2 diabetes.	Thank you for your comment.
Academy of Nutrition and Dietetics	Results – KQ4	*Treatment for Recurrent Gestational Diabetes Mellitus* The Academy evaluated five studies to investigate the relationship between nutrition interventions and the recurrence of gestational diabetes mellitus in women with a history of GDM. Studies reporting recurrence of GDM show a prevalence ranging from 30% to 65%. For women with GDM who are overweight/obese or with above-recommended weight gain during pregnancy, the Academy recommends that RDs should advise weight loss after delivery, which includes a combination of diet modification and physical activity. Research indicates that the risks of recurrent GDM or development of type 2 diabetes can be reduced with weight loss.	Thank you for your comment.

Commentator & Affiliation	Section	Comment	Response
Academy of Nutrition and Dietetics	Results – KQ5	The Academy notes that the Evidence Synthesis recognizes that the "evidence supports benefits of treating mild GDM with no evidence of harm… [and that] RCTs of GDM treatment show no evidence of harm related to treating GDM, other than an increased demand for services." However, the Academy has identified several risks of harms related to treatment for GDM: • Physical activity may cause hypoglycemia in women with gestational diabetes mellitus (GDM) using pharmacological therapy; • Contraindications to exercise during pregnancy may include but are not limited to: pregnancy-induced hypertension, premature rupture of membranes, intrauterine growth retardation, preterm labor or history of preterm labor, incompetent cervix/cervical cerclage, and persistent second or third trimester bleeding; • Frequent glucose self-monitoring may cause pain and discomfort; and • Use of pharmacological therapy to control blood glucose levels may result in hypoglycemia.	Thank you for your comment.
American Diabetes Association	Results – KQ3	We are unclear why this study, which compares outcomes of women with GDM by International Association of Diabetes and Pregnancy Study Groups (IADSPG) criteria but not World Health Organization (WHO) criteria, was not included in the evidence review for key question 3: *O'Sullivan EP et al. Atlantic Diabetes in Pregnancy (DIP): the prevalence and outcomes of gestation diabetes mellitus using new diagnostic criteria. Diabetologia 2011;54:1670-1675*	This study was identified by our literature search and later excluded from analysis due to the groups analyzed. The GDM group in the study encompassed women diagnosed with GDM by *both* WHO and IADPSG criteria. Women diagnosed by WHO criteria received treatment; those diagnosed with IADPSG GDM did not. The Normal glucose tolerance group in the study did not receive treatment. Because the GDM group had treated women, it did not meet our inclusion criteria regarding no treatment.

Commentator & Affiliation	Section	Comment	Response
American College of Obstetricians and Gynecologists	Results – KQ4, 5	The report looks at the following Key Questions: KQ4: "Does treatment modify the health outcomes of mothers who meet various criteria for GDM and offspring?" and KQ5: "What are the harms of treating GDM and do they vary by diagnostic approach?" The College suggests that data on cesarean delivery be added to more adequately address these questions. Currently, there is no discussion of the effect of treatment versus no treatment of GDM on the rate of cesarean section. At least two high-quality randomized controlled trials of treatment of lesser degrees of glucose intolerance (mild gestational diabetes) have reported the rate of cesarean section.	KQ4 addresses cesarean delivery in relation to treatment/no treatment, and both key questions utilize data from the Crowther and Landon trials.
American College of Obstetricians and Gynecologists	Results – KQ5	Under the discussion of Key Question 5, the report states that there was "no evidence for some of the outcomes stipulated in the protocol including costs, resource allocation, and healthcare system issues." Both the Crowther 2005 and Landon 2009 studies examined cesarean sections rates, rates of induction of labor, and number of clinic visits (prenatal or other). To the extent that an unnecessary cesarean section is harm, and that cesarean section rates are a highly important outcome, this should be reported or discussed in Key Question 5. Rates of induction of labor are also a tremendously important outcome or "health system issue" with great national visibility at this time. To the extent possible, this should be addressed in Key Question 5. Finally, the number of prenatal care visits is clearly a "health system issue" and one that looms large to US obstetricians and other providers trying to determine the impact of changing to and treating according to various proposed diagnostic criteria and the workload that follows. Both of the above studies addressed number of clinic visits.	We have changed the text to read "no evidence for some of the outcomes stipulated in the protocol including costs and resource allocation." We agree that cesarean section and induction of labor can be considered health systematic issues. We have discussed induction of labor and cesarean section under Key Question 5 as well as Key Question 4. We have reported on the number of clinic visits in Key Question 5.

Commentator & Affiliation	Section	Comment	Response
Hayley	Discussion	"Evidence supports benefits of treating mild GDM with no evidence of harm." It is important to be aware of the potential impact of this statement or further clarify 'no evidence of harm'. Although there may be no medical or physiological harm, treatment of mild (or false positive) GDM cases can rapidly increase our health care system and private insurance costs and place undue stress on our patients and physicians. This is especially true with the stricter diagnostic thresholds that are being released to identify GDM patients. A diagnosis of GDM can also place unnecessary stress on the pregnant woman since a diagnosis of GDM automatically categorizes her as 'high risk' thus increasing the need for additional testing (NSTs, BPPs, ultrasounds, dr appts, etc.)and monitoring. Encouraging treatment of mild GDM or potentially false positive GDM (based off of stricter diagnostic thresholds) places physicians in a situation where they must treat these low risk cases to the fullest to prevent medical malpractice lawsuits. This encourages over-testing, increased medical costs, more stress for both physicians and patients, and a higher workload for physicians who must closely monitor any patient categorized as high risk. It may be important to clarify the level of medical intervention associated with low-risk or mild GDM patients. Limiting the intervention to be more on education, diet modification, and a slight increase in monitoring needs to become standard for low-risk or mild GDM patients. Consideration of the full implications of expanding the number of newly "identified" mild GDM patients needs to be weighed when deciding on the level of 'harm' and determining the diagnostic thresholds for GDM.	We have added to our future research needs in order to address these comments. Specifically, we have indicated that future research needs to examine the long-term impact of a GDM label.
American Diabetes Association	Conclusions	We agree with the conclusions from the evidence, and we hope that these conclusions will inform the ultimate recommendation for GDM screening. We are hopeful that the pending USPSTF recommendation for GDM will align with the Task Force's expressed ambition to have fewer "I" statements and provide greater guidance to the nation's health care providers.	Thank you for your comment.

Commentator & Affiliation	Section	Comment	Response
Anonymous	References	I believe that there are key resources missing from this study, including (but not limited to) the following: Dabelea et al. (2000). Intrauterine Exposure to Diabetes Conveys Risks for Type 2 diabetes and obesity. Diabetes, 49, 2208-2211. Dabelea et al. (2005). Increasing prevalence of gestational diabetes mellitus (GDM) over time and by birth cohort: Kaiser Permanente of Colorado GDM Screening Program. Diabetes Care, 28(3) 579-584. Dabelea et al. (2008). Association of intrauterine exposure to maternal diabetes and obesity with type 2 diabetes in youth. Diabetes Care, 31(7), 1422-1426. Metzger, Boyd (2007). Long-term Outcomes in Mothers Diagnosed With Gestational Diabetes Mellitus and Their Offspring. Clinical Obstetrics and Gynecology, 50 (4), 972-979. Nolan, Christopher (2011). Controversies in gestational diabetes. Best Practice & Research Clinical Obstetrics & Gynecology, 25 (1), 239-244. Ogonowski, J. and Miazgowski, T. (2009). The prevalence of 6 weeks postpartum abnormal glucose tolerance in Caucasian women with gestational diabetes. Diabetes Research and Clinical Practice, 84(3), 239-244. Poston, Lucilla (2011). Intergenerational transmission of insulin resistance and type 2 diabetes. Progress in Biophysics and Molecular Biology, 106, 315-322. Rivas et al. (2010). Awareness of risk factors for type 2 diabetes in women with current and former gestational diabetes mellitus-implications for future primary diabetes prevention. Diabetes & Metabolic Syndrome: Clinical Research & Reviews, 89-94.	We have reviewed this list of studies provided by this reviewer and determined that they do not meet the inclusion criteria for any of the key questions in this report.