Evidence Report/Technology Assessment Disposition of Comments Report

Research Review Title: Smoking Cessation Interventions in Pregnancy and Postpartum Care
Draft review available for public comment from June 20, 2013 to July 17, 2013.

Research Review Citation: Likis FE, Andrews JC, Fonnesbeck CJ, Hartmann KE, Jerome RN, Potter SA, Surawicz TS, McPheeters ML. Smoking Cessation Interventions in Pregnancy and Postpartum Care. Evidence Report/Technology Assessment No. 214. (Prepared by the Vanderbilt Evidence-based Practice Center under Contract No. 290-2007-10065-I.) AHRQ Publication No. 14-E001-EF. Rockville, MD: Agency for Healthcare Research and Quality; February 2014. 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 evidence report/technology assessment 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 evidence report/technology assessment.

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://www.effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1871
Published Online: February 25, 2014

Commentator & Affiliation	Section	Comment	Response
Reviewer 6	Abstract	Objective should state 'reviewed evidence about smoking cessation interventions in pregnant…etc. It would also be helpful if the objective could state that it reviewed effectiveness and safety of the interventions for women and their infants.	We have revised the objective to include "interventions." We did not add the additional suggested text due to word count limitations.
Reviewer 4 (TEP)	Executive Summary	Background ES1: rates are presented as accurate/valid. The rates presented are significantly biased, and invalid. They are ONLY derived from CDC and each state (PRAMS/Post-Partum Surveys). Multiple SAMHSA reports based on annual, representative household face to face surveys each year from 1990 to 2010 confirm a much higher self-reported rate than the CDC rates. Although all rates are based on patient reports, the Report totally ignores the very large "Non-Disclosure Rates" confirmed at entry and during care (10%) by multiple publications, (e.g., range= 25%/ Windsor, et al 1993 to 48%/ CDC-SCIP/Kendrick, et al 1995).	We have inserted a statement in the Background section of the Executive Summary noting the range of nondisclosure rates from selected publications to reflect the fact that self-reported smoking and cessation rates vary and are underestimates. In response to a comment below, we have also added a section to the Introduction Chapter of the Main Report "Measurement of Tobacco Exposure During Pregnancy" which includes information about nondisclosure rates.
Reviewer 10 (TEP)	Executive Summary	Page 10 - ES-2 line 45: Question 1- Quitting smoking while pregnant is very beneficial to both the fetus and the mother.	We believe the reviewer is asking us to change "may be beneficial" to "is very beneficial." However, the sentence in question refers to the potential benefits of smoking cessation interventions not of smoking cessation itself. Therefore, the text is correct as written and was not revised.
Reviewer 10 (TEP)	Executive Summary	Page 9 - ES-1 line 16: "complications for both the pregnant woman and her fetus" this should also include extended family.	We have revised the sentence to include other household members.
Reviewer 10 (TEP)	Executive Summary	Also line 19: "Health risks to the fetus include low birth weight..." should also include chance of asthma, chance of behavioral problems, chance of SIDS.	There are a number of other health risks that could be added. We state the risks "include" (rather than "are") to indicate this list is not exhaustive.
Reviewer 10 (TEP)	Executive Summary	Also line 21: "...counseling, self-help materials..." should include motivational interviewing.	We considered motivational interviewing to be a counseling approach.

Source: http://www.effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1871
Published Online: February 25, 2014

Commentator & Affiliation	Section	Comment	Response
Reviewer 10 (TEP)	Executive Summary	Page 9 ES-1 line 50: “biochemically validated reports” cotinine or carbon monoxide? Cotinine is the gold standard. Page 10 - ES-2 line 23: Outcome-carbon monoxide or cotinine? (saliva in urine)	We included studies that used cotinine, carbon monoxide, or thiocyanate validation. Information about methods of biochemical validation has been added to the Introduction.

Commentator & Affiliation	Section	Comment	Response
Reviewer 1 (TEP)	Introduction	Well-done. Quite evidence based and consistent with my understanding of the state of knowledge in this area.	Thank you for your comments.
Reviewer 2 (TEP)	Introduction	Consistent and supports current literature.	We appreciate the comment.
Reviewer 4 (TEP)	Introduction	Background, ES1: The statement in the Report, "Self reports may lead to an overestimation of cessation rates during pregnancy" is obfuscation to an nth degree. The correct words are: "will lead to" not "may".	We have revised the sentence to read: "Furthermore, self-report leads to an overestimation of smoking cessation in pregnancy."
Reviewer 4 (TEP)	Introduction	Background: The Report should read: "Multiple studies of large, representative samples of pregnant at the onset, during after pregnancy have confirmed by urine, saliva, or carbon monoxide tests very high levels of patient non-disclosure. Large percentages of patients report on entry, during, and after pregnancy that they are not smokers when they are smoking." The Report must add multiple references here. Citing only the Boyd, Windsor, et al, 1998 in this section represents a very serious flaw in the Report that must be corrected.	We have elaborated by including additional citations and describing the significance of this issue. Specifically, we inserted the following text and cited seven references: "Multiple studies of representative samples of pregnant women using biochemical measurements have confirmed high levels of patient non-disclosure."
Reviewer 5	Introduction	The introduction is fine.	Thank you.
Reviewer 6	Introduction	The Key Questions are all important and appropriate.	Thank you for your comment.
Reviewer 6	Introduction	The rationale for including post-partum studies in this review, however, is unclear.	The topic request included postpartum women.
Reviewer 8 (TEP)	Introduction	The intro is well written and lays out what we are looking into. The report does not follow as well organized however.	We have attempted to make the organization of the report clear via introductory paragraphs for chapters and sections, frequent headings, and following EPC guidance for preparing the report. Unfortunately the reviewer does not provide specific critique or suggestions so we are unable to assess how to improve what s/he found lacking.
Reviewer 9 (TEP)	Introduction	The introduction clearly describes the clinical and policy context of the problem, the purpose of the report, and the key questions.	Thank you for your comments.
Reviewer 10 (TEP)	Introduction	Page 5 vi - line 34: I am unsure what is meant by "No significant harms"	We have revised to read, "No serious harms were identified…"

Commentator & Affiliation	Section	Comment	Response
Reviewer 11 (TEP)	Introduction	The Introduction was well-written and provided a concise overview of the state of the science. This review is to be commended for attempting to expand our knowledge base about infant and child outcomes.	Thank you for your comments.
Reviewer 11 (TEP)	Introduction	It was not entirely clear how the effect modifiers that were included in Key Question 5 were chosen. Further elaboration should be considered.	Effect modifiers were selected a priori from review of literature and in collaboration with the technical expert panel.
Reviewer 11 (TEP)	Introduction	I appreciated explicitly stating that public health initiatives or system-level initiatives were excluded.	Thank you.
Reviewer 11 (TEP)	Introduction	In my opinion, Figure 1 is vague. I think more explanation in the text would be helpful. Otherwise, the reader needs to give the figure some additional attention.	Figure 1 is the Analytic Framework which is presented using EPC methodologic conventions for a visual representation of the report Key Questions and outcomes of interest.
Reviewer 12	Introduction	The report could be clinically meaningful if there were more well-designed studies in the field. The target population is well-defined.	Thank you.
Reviewer 12	Introduction	Key Question 2 states that the report will examine child outcomes, but there are no follow-up results past 6 months postpartum. Follow-up should be at least a year after birth- and that is still termed an infant outcomes Child outcomes were not reported here so that term should be deleted from Key Question 2.	We included studies of pregnant or postpartum women who currently smoked or who had quit during the index pregnancy. Followup in these studies was short, and none reported child outcomes.

Commentator & Affiliation	Section	Comment	Response
Reviewer 1 (TEP)	Methods	The inclusion and exclusion criteria generally seem quite appropriate.	Thank you for your comments.
Reviewer 2 (TEP)	Methods	Inclusion and exclusion criteria are sufficient and conservative. Focusing on RCT and studies with biochemical validation provides the strongest evidence. Search strategies are clearly stated and outcome measures are appropriate. I'm not familiar enough with some of the statistical methods, so I will defer this to other researchers.	Thank you for your comments.
Reviewer 3 (TEP)	Methods	Inclusion criteria are appropriate and clear, a priori, reasoning is used to justify the use of only higher-quality studies to contribute to the research synthesis.	Thank you.
Reviewer 3 (TEP)	Methods	One issue which I think is a little confusing is that the intervention components which are investigated for KQ4 are not consistent with the KQ1 categories for which current research is qualitatively appraised. Consequently, KQ4 findings for using 'Incentives' are given prominence in the Executive Summary but there is no qualitative summary of trials which have investigated incentives or their participants, whereas there are for some other intervention components investigated for KQ 4 (e.g. NRT). As findings relating to incentives are given such prominence, perhaps the report should be structured such that trials investigating these described together in similar detail to the trials which describe intervention strategies in relation to KQ1?	There are two studies in which incentives were the primary intervention, and they are now classified as such in the "other" section for Key Question 1. The remaining studies that included incentives were multicomponent. The categories used to classify interventions for Key Question 1 and Key Question 4 are the same; however, Key Question 1 is organized by primary intervention while Key Question 4 examines the primary and secondary interventions in each study.
Reviewer 3 (TEP)	Methods	For KQ4, "Effectiveness of interventions",I think it could be useful to show how the descriptors of interventions compare with those used in other reviews (e.g. Cochrane).	We reviewed descriptors of interventions from other reviews in developing our definitions. Descriptors of interventions are similar across reviews, and differences primarily reflect minor variations in wording.
Reviewer 4 (TEP)	Methods	The Report, in addition to the current Assessments Form, should include as a criteria: Did the evaluation specify in the publication estimates of "Effect Size" and Statistical Power". If a study did not, this is a significant methodological error.	Estimates of effect size and/or statistical power were not required for inclusion in this review, and we did not extract this information from the included studies.

Commentator & Affiliation	Section	Comment	Response
Reviewer 4 (TEP)	Methods	In addition, the following study needs to be reviewed again. Dornelas reported a C Group Quit rate of 9.6%, E Group rate of 28.3%, and P = 0.015. Because of the small sample size, using these quit rates and a P= 0.01, I computed (P1 = 0.10 and P2 = 0.30) the sample size per E + C Group. A P = 0.01 requires > 71 patients per group at follow-up for a Two-Tailed Test. There is something wrong here.	We have rechecked the calculation. The probability of the variation (18.7%) between the two groups being due to chance was 0.01428 and 0.02856 for a one and two-tailed test, respectively.
Reviewer 4 (TEP)	Methods	This Report indicates that a study needed to have > 20 subject in the Experimental Group (E) and 20 in the Control/Comparison Group (C). Randomization and bio confirmation do not make RCT's of the studies included in the Report. Any study with approximately ≤ 50 patients per E + C Group is a Pilot or Feasibility evaluation. A review of AHRQ, 2000 and 2008, and all reports from Lumley, et al confirm that this criteria HAS NOT been applied in previous Meta-Analyses. The justification for this unacceptable sample size criteria in the Report makes no sense.	With the assistance of our technical experts, we have added rationale for the decision to consider inclusion of studies with 20 or more participants per arm to the Methods chapter. Briefly, we selected a minimum sample size of 20 in order to maximize our ability to describe the state of the current literature, while balancing the need to identify studies that could be used to assess treatment effectiveness. Furthermore, given the flexibility of the mixed-effects model employed in our meta-analysis, and the assessment of methodological rigor of individual studies (i.e., risk of risk of bias associated with randomization, selective outcome reporting, and attrition) we are comfortable with the sample size cutoff of 20 for this particular review. Of note, six of the 28 good and fair quality studies included in the review randomized fewer than 50 participants per arm. Of these, two studies evaluated relapse prevention in recent quitters. Therefore, four studies of pregnant current smokers with fewer than 50 participants per group contributed to the meta-analysis. The study by Heil (2008) was included in the systematic reviews by Greaves (2011) and Lumley (2009). The study by Hennrikus (2010) was included in the systematic review by Greaves (2010).

Commentator & Affiliation	Section	Comment	Response
Reviewer 4 (TEP)	Methods	All evaluation studies with less than < 50 patients in each group at follow-up (not at baseline), should be referenced, but should be excluded from the Meta-Analyses. While the reasons for this statement should be self-evident, the following confirm that using > 20 has no scientific base. Assume a Control Group Quit Rate of 5% (P1 = 0.05), the study would have to produce an E Group quit rate of 30% (P2 = 0.30), the E + C Group would have to have > 43 patients at the follow-up assessment to be statistically sig. (See CDC Epi-Info/Estimation of sample sizes + Fleiss). Having taught evaluation methods for 35 years at multiple Schools of Public Health, including a review of sample size estimation-calculation, these principles are taught to all of our MPH students. It is a basic principle of "Evaluation 101". The Report will lose credibility if it retains the > 20 subjects per condition in the Meta-Analysis.	Although it is commonly used, a minimum sample size of 50 has no intrinsic value as a cutoff. Rather the decision about sample size should be based on the purpose of the review and the methods used to synthesize the literature. One of the aims of a meta-analysis is to combine studies with low power to improve overall power. Hence, we do not use a standard cutoff for sample size of included studies, though we do acknowledge that very small studies are often flawed for other reasons.
Reviewer 4 (TEP)	Methods	How could 40-50 patients per group be representative of a clinical population of hundreds of patients from multiple sites? Inclusion of these studies with very small samples sizes cannot meet the criteria for "Internal Validity": valid for the eligible sites/populations . Thus, these studies also cannot meet the criteria for "External Validity". They are not "Representative" of the sites or diverse population at these sites. My Meta-Analysis in 2010 (See Attached) using a criteria of > 50 patients per group at the follow-up assessment indicates, at a minimum, that the following evaluations should be removed from the Meta-Analyses: Albrecht, 2006, Ondersma, 2012, Suplee, 2005, Hotham, 2006, Philips, 2012, Hennrikus, 2010, Price, 1991, and El-Mohandes-Windsor, 2012. Other studies may also not meet this criterion.	Please see our response above. Methodology guidance for EPC reviews does not stipulate a minimum sample size for inclusion of RCTs. Decisions regarding the type and size of studies are based on the availability and quality of existing evidence to address the review's Key Questions. The EPC prefers to include good quality, large, studies adequately powered to detect meaningful differences. Moreover, the EPC investigators consult a team of clinical and methodologic experts during protocol development to confirm that key decisions for eligibility are reasonable in light of the current state of the evidence. For this particular review the team deliberated on a minimum sample size and vetted the decision to use 20 as the cutoff with the technical expert panel.

Commentator & Affiliation	Section	Comment	Response
Reviewer 4 (TEP)	Methods	I strongly disagree with the rating of the studies as “Good”, “Fair”,”Poor”. All of my experience in performing a Meta-Analysis, indicates four levels of study methodological strength: Excellent, Very Good or Good, Fair, and Poor.	There are a variety of methods for assessing risk of bias. We used a three-level system for quality assessment corresponding to the Cochrane Risk of Bias categories of “Low”, “Unclear”, and “High”. This decision is supported by EPC Methodology (see: Viswanathan M, Ansari MT, Berkman ND, et al. Assessing the risk of bias of individual studies in systematic reviews of health care interventions. In: Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(12)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality; April 2012) and experience with numerous prior systematic reviews conducted by this EPC.
Reviewer 5	Methods	The inclusion and exclusion criteria are clearly stated in Table 2 and they are fine.	Thank you.
Reviewer 5	Methods	The search strategies are explicitly stated and logical, specifically the search strategy and search terms are stated clearly on page 7 of the report and reiterated elsewhere in the Methods section.	Thank you.
Reviewer 5	Methods	The definitions for the outcome measures are stated in the Methods section, self reported outcomes alone are not sufficient to be included in the report, smoking outcomes must be biochemically verified.	We only included studies that provided biochemically validated smoking cessation outcomes and have clarified this further in the text.
Reviewer 5	Methods	The statistical methods are OK.	Thank you.
Reviewer 6	Methods	Although the authors have attempted to justify their reasons, I think there are two main weaknesses in the inclusion and exclusion criteria. Firstly, only studies with validated smoking cessation are included. Although this is preferable and should be presented, it could have been compared with self-reported cessation. Valuable information could have been missed by not including these studies as although self-reported cessation may be exaggerated, in a randomised trial you would still have information on the difference between arms.	Validated smoking rates are more accurate; therefore, we chose, in consultation with our technical expert panel members only to include studies that reported biochemical validation. We agree assessing the difference between self-reported and validated smoking cessation outcomes is important, but that is beyond the scope of this review.

Commentator & Affiliation	Section	Comment	Response
Reviewer 6	Methods	Secondly, small studies were excluded, but if these had been included they could still have contributed to meta-analyses.	As there are no hard and fast rules for sample size of included studies, the investigative team selected a cutoff of 20 per group based on what was known of the available literature and a need to balance resources and the highest methodologic standards with the overall goal of addressing the Key Questions while limiting selection bias of included studies.
Reviewer 6	Methods	Page 4 Table A PICOTs: KQ2 – other important infant outcomes e.g. congenital abnormalities, postnatal death, disability or development, do not seem to have been included. Is there a reason for this?	We selected outcomes for which evidence supports an association with smoking. Neonatal death was included.
Reviewer 7 (TEP)	Methods	Yes - Although more studies would have been available if a biochemical confirmation of cessation had not been required. Also complete cessation is the ultimate goal but significant decrease is clinically useful. However that would have been an impossible consideration.	While not requiring biochemical validation would have increased the number of eligible studies, it would also have introduced the problem of the discrepancy between self-report and validation of cessation. We agree that decreasing smoking is clinically important; however, we chose to examine smoking cessation because it is the optimal outcome for maternal and infant health.
Reviewer 8 (TEP)	Methods	I believe all the definitions are appropriate, and outcome measures, as they are, are appropriate. By the report it seems there are not adequate studies available to answer the questions. But appropriate inclusion and exclusion criteria.	Thank you for your comments.
Reviewer 9 (TEP)	Methods	Inclusion/exclusion criteria, search strategies, definitions, and statistical methods are all appropriate for the project and expertly developed, described, and applied.	We appreciate your comments.
Reviewer 11 (TEP)	Methods	I couldn't determine whether an 'intent to treat' analysis was applied to all analyses. This could depend on whether the authors explicitly stated that they used this type of analyses, or whether the authors were able to extract this information from the paper, or contact the authors for this information. While this was considered in the 'risk of bias', it might be helpful to describe.	Studies that used intention-to-treat analyses were generally judged to have a low risk of bias for the incomplete outcome data domain. EPC Investigators do not routinely contact authors; therefore, if the study authors did not clearly report an intention-to-treat analysis or provide an explanation for how they handled data for participants lost-to-followup, the study quality assessment was downgraded for this domain.

Commentator & Affiliation	Section	Comment	Response
Reviewer 12	Methods	The inclusion criteria included only studies with biochemically validated smoking cessation outcomes. The authors should have justified this decision with reference to quality of measurement studies among this population. One reference they can use for pregnant women, although there may be others, is: Boyd, NR, Windsor RA, Perkins L and Lowe, JB: Quality of Measurement of Smoking status by Self-Report and Saliva Cotinine among Pregnant Women. Maternal and Child Health Journal Vol 2. No 2. pp77-83, 1998.	Information about the available methods of biochemical validation has been added and we cited the article by Boyd (1998).
Reviewer 12	Methods	The authors chose a sample size of 20 participants for each study arm as inclusion criteria. What is the power to test associations with this number? How was 20 chosen?	As there are no hard and fast rules for sample size of included studies in a systematic review/meta-analysis, the investigative team selected a cutoff of 20 per group based on what was known of the available literature and a need to balance resources and the highest methodologic standards with the overall goal of addressing the Key Questions while limiting selection bias of included studies.
Reviewer 12	Methods	On page 33, last line, please define “Frequentist confidence Intervals.”	We have deleted reference to “frequentist” and revised to “classical (non-Bayesian) confidence intervals.”
Reviewer 12	Methods	On page 38, 4 lines of text from bottom, rewrite “drilled down exploration” into something that is more clear and professional.	We have replaced “drilled down” with “detailed”.

Commentator & Affiliation	Section	Comment	Response
Reviewer 2 (TEP)	Results	The results are clearly stated. The tables and figures (abundant) help the reader review the studies and data. I did not see any areas that were overlooked.	Thank you for your comments.
Reviewer 3 (TEP)	Results	# 1 For studies investigating the use of NRT, it's not clear to me why the paper by Wisborg (2000) isn't included. This is a double blind, placebo-RCT and, I'd have thought would have 'good' quality and so could contribute to findings.	We excluded the study by Wisborg (2000) because biochemically validated abstinence was only reported by proportion; the number of participants included in the calculation of these percentages was not provided. Therefore, we could not ascertain whether the number of participants met our minimum sample size requirement. We contacted the author for clarification of the number of participants but did not receive a response.
Reviewer 3 (TEP)	Results	For multicomponent studies, Table 19 should record that the intervention delivered by Hegaard (2003), included an offer of NRT but his offer was only taken up by a minority of participants allocated to the intervention group.	Table 19 presents an overview of all of the study interventions that were offered. Individual use of specific interventions varied, as was the case with NRT in the study by Hegaard and colleagues. We did not extract data about the uptake of individual components.
Reviewer 3 (TEP)	Results	A mentioned above, it would be more transparent to have the KQ 4 the empirical literature on different components of interventions for described qualitatively together.	Table 22 presents an overview of the different intervention components across studies.
Reviewer 4 (TEP)	Results	I do not understand the criteria for the evaluations of "Educational Materials" included in Table 11.	Table 11 presents outcomes from all of the included studies that had educational materials as the primary intervention.
Reviewer 4 (TEP)	Results	Inclusion of Odersma, 2012, and rating it as "Good" is absurd.	The study by Ondersma and colleagues (2012) met the prespecified inclusion criteria. The quality was rated using the process described in the Methods section.

Commentator & Affiliation	Section	Comment	Response
Reviewer 4 (TEP)	Results	Report includes Moore, 2002, Ershoff, 1989, and Hjalmarson. (the "Windsor Guide was translated/adapted to Swedish with my permission) yet you do not include the three Windsor, et al evaluations: 1985-1993-2000.	All of the studies (Windsor et al. 1985 [Ref ID: 1332]; Windsor et al. 1993 [Ref ID: 1187]; Windsor et al. 2000 [Ref ID: 929]) are included in the review. Data from the studies reported in the 1985 and the 1993 publications are summarized in the multicomponent interventions section in Key Question 1 and contribute to the meta-analysis, with results reported in Key Question 4. The data from the 2000 publication contributes to Key Question 5. The 2000 publication is not included in Key Question 1 because a portion of the control group was nonconcurrent.
Reviewer 4 (TEP)	Results	The AHRQ 2000 and 2008 Guidelines cited the 1986 and 1993 SCRIPT evaluation as "Excellent". In addition, the priority scores (100 to 500) for Trial I was 122 (5th percentile) and Trial II it was 117 (0.6th Percentile).	This review provided a de novo, objective review of the literature using the methods described in the text. We stand by our assessment of this study as reported. We do not take priority scores into account in our assessment of the literature.
Reviewer 4 (TEP)	Results	I do not understand why the Windsor, et al, 2000, Journal of Ob/Gyn, 2000, pp 68-75 study was not reviewed in the Meta-Analysis. Considering the very large number of very weak studies reviewed in the Report, I believe that the methodological strength of the evaluation in Jnl. Ob/Gy, 2000 merits inclusion. It was the 1st "Effectiveness" evaluation to a representative cohort of Medicaid supported patients for a State-Wide (Alabama) prenatal care program. The SCRIPT Program was delivered by 28 regular prenatal care (RN/SW) staff to 549 patients, including a biochemical assessment at entry and during the 3rd trimester, as an integrated component of primary care? The evaluation covered year 1-2. The problem of Control Group contamination (20% provided SCRIPT Program) was confirmed in year 3-4-5.	The results from the study published by Windsor and colleagues in 2000 were not included in the meta-analysis because this study did not meet the criteria for an RCT due to the fact that a portion of the control group was nonconcurrent. The study data did contribute to Key Question 5.

Commentator & Affiliation	Section	Comment	Response
Reviewer 4 (TEP)	Results	I do not agree with your Meta-Analysis of the NRT literature. I have reviewed the Myung Meta-Analysis. The Myung, et al report is cited, but the Report does not identify the actual study included but excluded by the Cochrane/Lumley Meta-Analyses. There is insufficient evidence that NRT is effective in assisting a significantly higher percent of pregnant smokers to quit. If she is not a reviewer, I strongly recommend that you ask Cheryl Oncken, MD, U. of Connecticut to review this section. Note: Because it is a very small group of senior investigators (NRT), we have all had ongoing face to face/electronic communication.	We did not conduct a meta-analysis specifically of the NRT literature but rather conducted a random effects meta-analysis to quantify the relative impact of components of the interventions on smoking cessation. Differences in the findings of this review compared to previous reviews likely reflect the fact that the reviews differed in inclusion and exclusion criteria, meta-analysis methods, etc.
Reviewer 4 (TEP)	Results	I do understand a rating of "Good" for the El-Mohandes/Windsor NRT study as "Good" and a rating of "Fair" for the Oncken, 2008 study. The Oncken NRT evaluation was much stronger than our "Formative Evaluation"	The quality of all studies was assessed as described in the Methods Chapter. Please also see Appendices C – F for additional details related to quality assessment. We have reassessed the risk of bias for the study published by El-Mohandes (2012) and have changed the quality rating to poor.
Reviewer 4 (TEP)	Results	I also strongly disagree with the recommendation about monetary incentives/Heit. This is ONLY one study with a very small sample size conducted in Vermont which has almost no racial/ethnic diversity. Thus, it has questionable Internal Validity and does not reflect external validity. The Report also ignores the ethical-practical-cost/patient of paying a smoker to quit. Should we pay (bribe) an individual to come to prenatal care early, to keep each visit, to change their diet, to stop drinking/etc., to lose/gain weight, to take vitamins, etc. I do not believe that primary care agencies are going to or should spend an additional $200-$300 per smokers. This Report needs to discuss these issues.	Four of the good or fair studies in this review included an incentive intervention: Ondersma et al, 2012; Walsh et al, 1997; Heit et al, 2008; and Donatelle et al, 2000 (see Table 22), and the odds ratio for the relative impact of incentives in the meta-analysis was 3.23 (Bayesian credible interval, 1.98-4.59). As noted in Table 27, all smoking cessation interventions require human, financial, and/or other types of resources. End users choosing which interventions to use will have to determine how to best allot their resources in terms of what they have available as well as how likely specific interventions are to be effective.

Commentator & Affiliation	Section	Comment	Response
Reviewer 4 (TEP)	Results	Effectiveness of Intervention Components (KQ4), p. 52: This part of the Report needs to be expanded to include a new section: "Rating Fidelity of Patient Assessment + Treatment Delivery: Process Evaluation". In addition to adding a Section on Measurement, It is equally important to define who (What percent of eligibles) participated and what core assessment and treatment procedures they received. The 29 "Good + Fair" studies, excluding the deleted evaluations with small sample sizes, would be the candidates for this Section. Windsor, et al, 2000, pp 29-35 (851/p 169) provides a detailed description, with examples, of how to "Rate" program "Fidelity." It was a major paper in the special issue of Tobacco Control, and was written for use to conduct a Process Evaluation of the 11 funded projects by the Robert Wood Johnson Foundation at the National Program Office (Robert Goldenberg, MD MPH, PI) in the Dept of Ob/Gyn, at the UAB Med Sch. I was the lead scientific advisor (40% FTE/ 1994-2000). The discussion in the current report is very weak about rating the "Fidelity of Treatment Delivery" The two paragraphs describe what should be included in the new section. Each study needs to be reviewed for the following "Process Evaluation" information. Note: Very few evaluations will report this information. The evaluation/documentation of the actual delivery of all core program assessment and treatment components by each study needs to be confirmed. Failure to treat needs to be examined. A detailed description of the SCRIPT Trial III process evaluation is presented on pp 415-416: #96, Health Education and Behavior, 2011. It documents for each provider and all providers how well they provided each/all "Procedures (P)". It is very important for this Report to present to the reader what the P1 was for each study in this section. This rate defines the denominator for each study and defines the degree of "Selection Bias".	We agree that fidelity of patient assessment and treatment is an important consideration in the effectiveness of interventions; however, evaluation of fidelity was not included in this review and thus cannot be reported. We have noted the lack of consistent reporting of fidelity as a limitation of the evidence base, and assessing intervention fidelity is noted as a future research need (see the Discussion section).

Commentator & Affiliation	Section	Comment	Response
Reviewer 4 (TEP) (continued)	Results	A process evaluation in SCRIPT Trial III confirmed that 6514 patients were screened…: 77% of the eligible smokers agreed to participate. Because our target participation % was 80% (patients decide to receive care). P1 = 77%/80% = 0.96. A Process Evaluation needs to answer: What number/% of patients who agreed to participate had a baseline assessment (P2) and a sample of saliva, urine, or CO collected (P3). In SCRIPT Trial III, the P2 and P3 baseline rates were 100% + 99%: P2 = 1.00 and P3 = 0.99. The three component, Video (P4), Guide (P5), and Counseling (P6), SCRIPT Program was delivered at an almost perfect level: P4 = 0.95, P5 = 0.99 and P6 = 0.97. The process evaluation all confirmed that the regular staff also completed 82% (P7 = 0.82) of follow-up assessments and 72% (P8 = 0.72) of saliva collection. The aggregation of all of these rates for the 28 RN's/SW's at the 8 clinics for a three year period for 1189 patients was a PII (Program Implementation Index) = 0.925. These process evaluation methods and data specific documentation that the SCRIPT Program was/can be delivered to a large Medicaid supported population with a very high level of "Fidelity". This is how a cessation program for pregnant smokers should document that the staff delivered how much of what to whom.	We agree that fidelity of patient assessment and treatment is an important consideration in the effectiveness of interventions; however, evaluation of fidelity was not included in this review and thus cannot be reported. We have noted the lack of consistent reporting of fidelity as a limitation of the evidence base, and assessing intervention fidelity is noted as a future research need (see the Discussion section).
Reviewer 5	Results	The amount of information and detail presented in the results section is quite extensive. The results section begins with an introduction, the goes through the results of the literature searches and the description of the studies included and then reports results based on each of the five key questions: intervention outcomes for pregnant and postpartum women; intervention effects on infant outcomes, intervention harms for pregnant and post partum women, effectiveness of intervention components; and effect of patient characteristics on effectiveness.	Thank you.
Reviewer 5	Results	The studies are described.	Thank you.

Commentator & Affiliation	Section	Comment	Response
Reviewer 5	Results	the key messages are identified as “key points” in each of the key question sections of the results section (five different “key point” in each of the key questions for the study.	Thank you.
Reviewer 5	Results	the figures, tables and appendices are both adequate and descriptive.	Thank you.
Reviewer 6	Results	It is stated that Wisborg et al 2000 RCT of nicotine patches for pregnant smokers has been excluded because it did not have validated cessation, but this paper (including the abstract) clearly states that they did validate cessation and so should be included.	We excluded the study by Wisborg (2000) because biochemically validated abstinence was only reported by proportion; the number of participants included in the calculation of these percentages was not provided. Therefore, we could not ascertain whether the number of participants met our minimum sample size requirement. We contacted the author for clarification of the number of participants but did not receive a response.
Reviewer 6	Results	The quality of the El-Mohandes et al 2012 study was classified as good. However, this study was not placebo controlled and although researchers may have been blind the participants would not have been blind to their treatment allocation; therefore the risk of bias for blinding of patients/personnel should not be classed as low.	We have reassessed the risk of bias for this study. We asked two team members to independently assess the study using the risk of bias form and criteria. The reviewers agreed that participants in the control group did not receive a placebo patch and thus were not blinded to their group assignment. Knowledge of treatment assignment was likely to influence the outcome. Therefore, we have changed the quality rating of this study to poor and have revised the report descriptions, counts, tables, and analyses accordingly.
Reviewer 6	Results	I found it unclear from the report how studies had ‘incomplete outcome reporting’ assessed. This is important as studies were classed as fair rather than good if they just had one ‘unclear’ score. For example, the Oncken et al 2008 study was only classified as fair.	Incomplete outcome reporting is assessed by determining the extent to which a study described the completeness of outcome data for each main outcome, including attrition and exclusions from the analysis; reported attrition and exclusions; and identified reasons for attrition/exclusions. More detail is provided in Appendixes C and D.
Reviewer 6	Results	Page 31 and Page 53: Table 12. Effect column: Coleman et al study found a doubling of quit rates in NRT group at 1 month, but no difference at delivery. This study also included a quit guide (available in online appendix to paper), but this is not mentioned here or in Table 22.	Thank you for pointing out these inaccuracies. We have added the information about the one month results to Table 12 and the Detailed Synthesis for NRT in Key Question 1. We have noted the use of the quit guide in Tables 12 and 22 and included this study in the meta-analysis.

Commentator & Affiliation	Section	Comment	Response
Reviewer 6	Results	Overall, when presenting cessation data it is unclear if the studies are reporting prolonged cessation or point prevalence. Prolonged cessation is more important, but rates will be much lower than a point prevalence measure.	We reported point prevalence because biochemical validation measures cannot assess prolonged cessation.
Reviewer 6	Results	Tables were a strength of the report and were very clear.	Thank you.
Reviewer 7 (TEP)	Results	The report is very detailed, very clear (for a topic that is anything but clear) and almost mind numbing for the careful detailed explicit explanation of the interpretation.	We appreciate your comments.
Reviewer 8 (TEP)	Results	At times too detailed. Repetitive presentation of the studies. But the tables were helpful. I think more simplified data would have been better in the body of the report and leave the specifics to the appendix.	We appreciate your comments.
Reviewer 9 (TEP)	Results	The results are clearly described, the key messages explicit, the tables useful. I am unaware of any inappropriate inclusions or exclusions, and, as a co-author of one of the studies receiving a poor quality rating (primarily because of lack of blinding), I believe the quality ratings were fairly and appropriately applied.	Thank you for your comments.
Reviewer1 (TEP)	Results	Results section generally appears to be solid.	Thank you.

Commentator & Affiliation	Section	Comment	Response
Reviewer1 (TEP)	Results	The Donatelle et al. (2000) trial is currently included in the category of "multi-component" interventions rather than "incentives". I suspect that decision is related to the intervention involving financial incentives and social support. However, the way the social support component worked was if a pregnant/newly postpartum smoker in the incentives conditions abstained from smoking both she and her social support partner received incentives. That is, the social support component was part of the incentives intervention. The Donatelle et al. trial was included as an incentives intervention in a recent review that my colleagues and I co-authored on the use of incentives among pregnant and newly postpartum smokers (Higgins et al., Prev Med. 2012 Nov;55 Suppl:S33-40. Epub 2011 Dec 27. Review. PMID: 22227223) and in the highly influential Lumley et al (2009) meta-analysis.I recommend consideration be given to relocating this trial in the "incentives" category.	In light of this comment, the team reevaluated whether this study should be categorized as an incentive or multicomponent study. The study met the team's definition of multicomponent for the purposes of this review. One investigator applied consistent criteria to all studies to identify and categorize individual intervention components, summarized in Table 19. As noted there, this study included five distinct components in the experimental arm, two of which were not part of the treatment delivered to the control group. When studies included more than one primary intervention, we made the methodological decision to group them as multicomponent based on the rationale that we could not tease out the effects of the individual interventions. The meta-analysis provides a complementary strategy by analyzing as much as possible the contribution of intervention components. We appreciate this reviewer's approach as well; by looking at the challenge of smoking cessation in this population through different approaches, we expect that it will likely become easier to triangulate the utility of the varied strategies explored in the literature.
Reviewer 11 (TEP)	Results	On page 41 of 388, the words 'drilled down' and 'drivers' are used. I would suggest replacing this jargon with more precise text.	We have replaced "drilled down" with "detailed" and replaced "be drivers of" with "promote".
Reviewer 11 (TEP)	Results	In Table 5, 'NR' is not included in the Abbreviations	We have added "NR=not reported" to the list of abbreviations.
Reviewer 11 (TEP)	Results	Table 6 would benefit from a clearer definition of "Relapse Prevention %". It isn't entirely clear what 81% relapse prevention means?	Relapse prevention indicates the woman has not resumed smoking and is synonymous with continued cessation. We have added a definition to the table footnotes.

Commentator & Affiliation	Section	Comment	Response
Reviewer 12	Results	The use of Good, Fair and Poor studies in the analysis is confusing and inconsistent. For example, for effectiveness, strength of evidence used good and fair studies, while infant outcomes and harms included poor quality studies in the strength of evidence assessment. I realize that this was done because of the number of studies in each area, but it makes no sense. It is not recommended to include poor quality studies in the review conclusions. Should the poor rating be reconsidered on some of these studies? Otherwise, EBR's do not usually use poor quality studies to come to conclusions.	Harms data from interventions are not typically well documented. We seek the best set of evidence, and for effectiveness, that is often RCTs. However, maintaining this standard of evidence for all types of outcomes (e.g., harms, infant outcomes in this review) would result in little to no evidence available for the report. That said, it is important for the end user of the report to know the quality of the literature from which the evidence arises.
Reviewer 12	Results	In some cases, statistically significant results were ignored because they were not "Clinically Relevant." I think it should be left to the reader to make that determination (page 14 under KQ2 paragraph).	The location provided in the comment does not match its content. We have searched the document and cannot locate the phrase "clinically relevant" so are unable to ascertain the specific content the reviewer thinks should be revised.
Reviewer 12	Results	In Table C of the Executive Summary, last column, the terms Low for no effect, low for effect are confusing and need to be rewritten or deleted.	As noted in the Methods section, we considered the estimate of the effect to be positive if the posterior probabilities based on the Bayesian credible intervals suggested greater than 80% likelihood that the true effect was greater than the null. To clarify, we have added a footnote to the Strength of Evidence Tables in the Executive Summary (Table E) and in the Main Report (Table 28).

Commentator & Affiliation	Section	Comment	Response
Reviewer 1 (TEP)	Discussion	Implications are clearly stated.	Thank you.
Reviewer 2 (TEP)	Discussion	I would consider this research easily translated to clinicians. The researchers did a good job stating the limitations. These findings have implications for clinicians and researchers.	We appreciate the comments.
Reviewer 3 (TEP)	Discussion	KQ2 findings are different from those of at least one other authoritative review (i.e. Cochrane) in finding no impact of cessation interventions on birth outcomes. Presumably this has arisen because only higher quality studies were included in the current review. The report would benefit from some speculation as to why these findings differ between reviews.	We added birth outcomes results from the Cochrane review and the following text to the section "Findings in Relationship to What is Already Known": "Findings regarding birth outcomes, including birth weight and preterm birth, are inconsistent across reviews. These differences may reflect the fact that reviews differed in inclusion and exclusion criteria and thus the studies that were included."
Reviewer 3 (TEP)	Discussion	KQ2 findings: generally more focused discussion of how findings differ from those of other authoritative reviews may also be beneficial.	We have added more detail as noted in the previous response.
Reviewer 3 (TEP)	Discussion	On page 69: "Overall, the findings from existing systematic reviews suggest that NRT, behavioral and educational cessation strategies, and multicomponent interventions may be beneficial to women who smoke in pregnancy or the postpartum period" . This statement is not correct with respect to the Cochrane review for pharmaceutical interventions (ref #34).	Findings from existing reviews regarding NRT have been conflicting thus it is accurate to say "may be beneficial," which acknowledges the fact that benefit is uncertain.

Commentator & Affiliation	Section	Comment	Response
Reviewer 5	Discussion	The findings are clearly stated, but I do have some concerns with the significant emphasis placed on "incentives" in this report. There are a number of major issues and concerns that have been raised in the literature with incentives and I worry a bit that the authors are catering to the current trend that the wind is blowing in the direction of "nudges" and other types of incentive programs, the effects of which it is well documented attenuate over time. In fact there is a real danger here in emphasizing the incentives approach and it is not clear that the analysis and review are sufficiently done or done by the right experts to justify these implications. This is a problematic part of the report that needs to be addressed before publication.	In our meta-analysis, incentives demonstrated the greatest effect among the components studied. This finding is based on our review of the evidence and not any current trend. We recognize that the use of incentives is controversial and have noted that in the report. Our role is simply to present the science, not to engage in that debate.
Reviewer 5	Discussion	Limitations: Not adequately in terms of the major limitations of incentives and the general approach of incentives/nudges over and in contrast to other approaches. A much more critical view needs to be included as we know that these "effects" are not long lasting, attenuate over time, and may even be misleading.	Followup beyond birth was rare thus we cannot compare how relapse rates vary by intervention. As noted in the Applicability section, little is known about durability of effects. We have added a comment about this to the limitations section.
Reviewer 5	Discussion	I think the whole criticism of incentives and these economic and monetary approaches, that is evidence based, has been left out of this report and it should not be published without a more balanced assessment.	The focus of this report is the effectiveness of the interventions. Certainly there are many considerations in choosing an intervention beyond effectiveness, some of which are noted in Table 27. Decisions about appropriateness and applicability of specific interventions will be made by end users, and we hope that we have provided adequate scientific information for them to make those decisions specifically for their populations.

Commentator & Affiliation	Section	Comment	Response
Reviewer 5	Discussion	Future research: The scope of the report was quite limited, which made it "cleaner" for the authors, but a number of different types of studies that are related to the study/target population and to the five key questions were not included, no economic evaluations were included, no second hand smoke, alternative study designs, self reported outcomes, etc.	The investigators, with advice from our technical expert panel, chose to limit the included study designs and require biochemical validation to ensure a rigorous review. We agree that a cost-effectiveness analysis would be worthwhile. However, as the reviewer also recognized, cost-effectiveness analysis is outside of the scope of this comparative effectiveness review. Given our limited review resource and timeline, we were not able to initiate de novo cost-effectiveness analysis and value of information analysis
Reviewer 7 (TEP)	Discussion	I anticipated that there would not be a single best (or a couple clear best) approach(es). This is an issue that has multiple facets to why an individual starts smoking, continues to smoke and what will motivate her to quit. This like many behavioral issues (obesity and alcohol abuse) are not easily approached and there will not be a single best answer which tends to dilute out the results to look like there is not clearly effective intervention.	We agree smoking and cessation are multifaceted and complex to address.
Reviewer 8 (TEP)	Discussion	Summary data shows that incentives work best, but combined efforts are effective. I am not sure that a lot of the information is not intuitively obvious. Saying that we are adding to the knowledge base because we include biochemical analysis data seems a stretch. The comment that incentives take less time and effort is intriguing though. Cost analysis might have been a good idea on all forms of counseling, incentives, etc. But again the studies were so short lived with no long term data. I would think the research that is needed is very broad.	Our meta-analysis and resource considerations for applicability are new contributions to the knowledge base. Cost effectiveness is not considered in EPC comparative effectiveness reviews.
Reviewer 9 (TEP)	Discussion	The discussion clearly lays out the implications of the major findings, the limitations of the evidence and the reviews, and the future research section presents a clear roadmap for moving forward.	Thank you.
Reviewer 9 (TEP)	Discussion	Table 27's discussion of resource implications of the different interventions is particularly helpful.	Thank you.

Commentator & Affiliation	Section	Comment	Response
Reviewer 9 (TEP)	Discussion	Although it is clearly beyond the scope of the report, some discussion of the potential use of the findings of this report for cost-effectiveness analyses of potential strategies for smoking cessation in pregnant and postpartum women is worthwhile. Although the lack of clear evidence in favor of one method or combination of methods prohibits a definitive analysis, such an analysis could be very useful for identifying strategies and/or combination of strategies which are optimal from a policy perspective. This is particularly true given the potentially controversial finding that incentives are among the most effective strategies. A cost-effectiveness analysis, combined with a value-of-information analysis, could be very useful for prioritizing among the many important evidence gaps identified by the report.	We agree that a cost-effectiveness analysis would be worthwhile, but the EPC program does not conduct cost-effectiveness analyses.
Reviewer 9 (TEP)	Discussion	Given that financial incentives seem to work, and that there may be some controversy about the concept of "paying people to do something they should be doing anyway", it might be worth a few lines about the use of financial incentives for other clinical or public health purposes.	We note that prior research in other fields suggests that modest incentives can be adequate to change behavior, and some references are provided.
Reviewer 11 (TEP)	Discussion	The implications of the major findings were clearly stated. The discussion of the effectiveness of NRT was important and received adequate attention. Clinically, this information is important and it was a strength that the limitations surrounding the NRT studies were adequately addressed. Future research topics are appropriate and translated into new research.	Thank you.
Reviewer 11 (TEP)	Discussion	On Page 96 of 31, the fourth paragraph mentions 'extremely heavy smokers'. This could be restated as 'highly dependent' smokers or those with 'extremely high levels of nicotine dependence'.Table 27 Component 'Groups' could be restated as 'Group counseling'.	We have revised the wording to "populations with extremely high levels of nicotine dependence" in the Executive Summary and Main Report. We use the term Groups to be inclusive of a variety of group approaches, not all of which may include counseling.

Commentator & Affiliation	Section	Comment	Response
Reviewer 12	Discussion	The Applicability section of the Executive Summary is confusing. Although incentives had the highest independent effect, the authors still suggest that a multicomponent intervention might work for providers.	While incentives had the highest independent effect, there were a number of other interventions that had a high likelihood of positive effects making a multicomponent intervention reasonable.
Reviewer 12	Discussion	More discussion should have been presented on the nature of the incentives and how they were given out.	The number of interventions and their variation across studies is huge thus we have provided an overview of studies. More detailed description of specific interventions in individual studies would make the report unwieldy. References are clearly identified so that interested readers can locate the original studies for further details.
Reviewer 12	Discussion	There is no good evidence on what other components should be picked since none had a left-sided confidence interval of 1.00 or over in Table C.	Of the nine interventions in Table E in the Executive Summary (previously Table C) and Table 28 in the Main Report, six demonstrated an effect (feedback, incentives, information, personal followup, quit guide, and NRT) while three did not (clinic reinforcement, peer support, and prescription to quit). That information, in conjunction with the resource considerations presented in Table 27, can be used in determining which interventions to include in a multicomponent model.
Reviewer 12	Discussion	Why aren't the 5 A's included as an intervention component in and of itself?	We grouped interventions into broad categories to keep the discussion meaningful and manageable. There were a number of counseling strategies as well as a wide variation in how well the type of counseling was described thus we combined these.

Commentator & Affiliation	Section	Comment	Response
Reviewer 1 (TEP)	General	No changes to recommend to organizational structure.	Thank you.
Reviewer 1 (TEP)	General	I found this report to clinically meaningful, and target population well defined; not so sure about the intended audience being defined. The key questions were appropriate and explicitly stated.	We appreciate your comments.
Reviewer 2 (TEP)	General	The report is clinically meaningful. The authors did a good job capturing details of effective interventions. The target population and key questions are clearly defined.	Thank you.
Reviewer 2 (TEP)	General	Conclusion can be used for practices (provider) decisions. The report is well organized.	Thank you for your feedback.
Reviewer 3 (TEP)	General	Given the complicated nature of the literature being reviewed and also the low quality of much literature in this area, the organisation and structure of the report is satisfactory and reviewers manage to relate conclusions clearly.	Thank you for your comments.
Reviewer 3 (TEP)	General	Generally, I think the authors have done an exemplary job; this literature is extremely complicated and they have made a strong attempt at summarising it.	Thank you.
Reviewer 3 (TEP)	General	Key questions are clearly defined and the target audience of the review (health care providers) is explicitly stated.	Thank you.
Reviewer 4 (TEP)	General	A Table that presents a synthesis of this literature should be added and is an essential reference for users of the Report.	The scope of this review was to focus on the treatment literature. A review of the literature on assessment of smoking status would be interesting and important, but it is not possible in the context of this project. We have added additional references supporting the fact that non-disclosure of smoking is common.
Reviewer 4 (TEP)	General	While I realize that an “Executive Summary” will be prepared, I strongly urge my colleagues at Vanderbilt to step back and address the issue of its length as a barriers to its dissemination and use. It is too long	The Executive Summary was prepared according to the AHRQ EPC template. Additional materials, including journal manuscripts, will be developed to aid in dissemination of the report findings.

Commentator & Affiliation	Section	Comment	Response
Reviewer 4 (TEP)	General	Because it builds on the AHRQ, 2000 and 2008 for two high-risk populations, It is essential to present a brief description of how the methods used to prepare this Report and the 2000 and 2008 Reports were comparable or different. A new section is needed that provides a synthesis of specific "new" knowledge since AHRQ, 2008.	Both the 2000 and 2008 reports are clinical practice guidelines from the Treating Tobacco Use and Dependence Guideline Panel. This report is a systematic review of the evidence, not a guideline. This review is intended to provide updated scientific evidence that guideline developers can use to produce guidelines similar to those described by the reviewer The process for conducting this systematic review was separate and independent from that used for prior guidelines.
Reviewer 4 (TEP)	General	Preface page iii. The time period of the Report should be defined (iii). I assume the cut-off for the review was December, 2012.	Thank you. We are required to use the AHRQ EPC template for the document preface material. The cutoff is included in the Methods. "Searches were executed between October 2012 and January 2013."
Reviewer 4 (TEP)	General	Preface page iv. In the Preface of the Report (iv), It should define for the reader the difference between "Effectiveness" and Comparative Effectiveness" Evaluation, using a concrete "Best Practices" example from among the RCT's reviewed.	Thank you. We are required to use the AHRQ EPC template for the document preface material.
Reviewer 4 (TEP)	General	In addition, I recommend that you consider significantly reducing the number of references. Many studies cited, including the evaluation studies, are not rigorous (See latter sections of this critique).	We included studies that met prespecified criteria for inclusion in the report. We address critiques of individual studies in subsequent comments.
Reviewer 4 (TEP)	General	This Report should be critical of multiple federal agencies for ignoring the large body of scientific evidence related to very poor measurement. High blood pressure rates, H1AC, HDL/LDL/Trig's, etc rates are all based on valid methods to confirm these clinical conditions. You would not ask a patient what her blood pressure was, or what their hemoglobin A 1C was. Past AHRQ Guidelines have ignored this major measurement problem, presenting one or two sentences about "Non-Disclosure". Poor measurement equals non-science. It needs to be directly addressed by this Report.	It is beyond the scope of our report to comment on problems with reporting by agencies; however, we acknowledge measurement of smoking status is a complex issue.
Reviewer 4 (TEP)	General	A Table that presents a synthesis of this literature should be added and is an essential reference for users of the Report.	We have added an evidence map for smoking cessation and relapse prevention to the Executive Summary The report contains multiple tables that synthesize this literature, including several that provide a broad overview (e.g., Tables 5, 6, 22-30).

Commentator & Affiliation	Section	Comment	Response
Reviewer 4 (TEP)	General	Because of the self-evident salience of the major measurement issue, this Report will be seriously deficient if it does not add a separate, new section: “Measurement of Tobacco Exposure During and After Pregnancy.”	We have added a subheading to the Introduction and elaborated on the issues related to measurement of smoking status in this patient population.
Reviewer 4 (TEP)	General	If this Report is the most current scientific statement about this area of research and practice, this measurement problem must be identified in this Section. I expect our Expert Panel + Vanderbilt Panel to review the measurement science as thoroughly as it has the treatment science. The two issues are inseparable.	The scope of this review was to focus on the treatment literature. A separate review of measurement issues would be interesting and important, but it is not possible in the context of this project. We have commented on issues related to measurement of smoking status in the Introduction and Discussion sections.
Reviewer 5	General	The report is clinically meaningful in the sense that it is trying to sort out the evidence base for smoking cessation and relapse prevention programs during the pregnancy and post partum period	Thank you for your comments.
Reviewer 5	General	The target population and audience are explicitly defined. The target of the search is smoking cessation and relapse prevention interventions during pregnancy and in the post partum period. The audience for the report is anyone who is interested in the evidence base on these interventions, at least the evidence based derived primarily from RCTs. Health care providers, however, are the main audience for the report. The interventions included were primarily individually based, as opposed to population or health systems oriented.	Thank you for taking the time to review and provide comments.
Reviewer 5	General	The key questions are appropriate and explicitly stated. The main purpose is to conduct a systematic review of the literature on smoking cessation and relapse prevention interventions during pregnancy and in the post partum period. Smoking outcomes were limited to those that were biochemically verified. The report did not include smoking reduction. The report used the PICOTS (population, intervention, comparator(s), outcome,timing and setting) structure.	Thank you for your comments.
Reviewer 5	General	The report is well structured and organized.	Thank you.
Reviewer 5	General	Main points are clearly presented.	Thank you.

Commentator & Affiliation	Section	Comment	Response
Reviewer 5	General	Main conclusion about incentives is not presented in a balanced enough way, it comes off as “the” most “effective” approach and just isn’t clear that the data support such a stand out conclusion, I would be very concerned about this.	In our meta-analysis, incentives demonstrated the greatest effect among the components studied. However, we noted multiple other interventions also demonstrated effect. Decisions about appropriateness and applicability of specific interventions will be made by end users, and we hope that we have provided adequate scientific information for them to make those decisions specifically for their populations.
Reviewer 6	General	This is a very long, comprehensive and interesting report and the authors should be applauded for the enormous effort that they have put into this piece of work.	Thank you.
Reviewer 6	General	This is an extremely lengthy review and therefore it is fairly difficult to find your way around. As the authors have included such a variety of studies, have attempted to answer a number of different questions and include a very wide range of outcomes it has perhaps not surprisingly become a quite confusing article to read and digest.	The broad scope of this review has resulted in a lengthy report. Additional materials, including journal manuscripts, will be developed to aid in dissemination of the report findings.
Reviewer 6	General	The text is quite ‘solid’ and it would be helpful to break it up more by using bullet points or numbering. This would be useful for all sections, particularly when describing recommendations.	We have used bullet points judiciously.
Reviewer 7 (TEP)	General	Since it did not find a clear best approach (which I did not expect personally) I worry that this may decrease the willingness of insurance/government to reimburse for smoking cessation activities. Other than that it stresses the need to continue to explore options. I feel Table 27 is a factual, accurate summary of what to do clinically given this review	Thank you for your comments.
Reviewer 7 (TEP)	General	Yes Again I was getting very depressed (but not surprised) by the findings. I found that Table 27 was a very nice summary of how this work impacts clinical activities	Your comments are appreciated.
Reviewer 8 (TEP)	General	Yes it is meaningful. Key questions are clear, population is defined.	Thank you.

Commentator & Affiliation	Section	Comment	Response
Reviewer 8 (TEP)	General	Tables are helpful to make the report more organized but it is quite long. I think we need stronger statements at the end of what we need. Those of us who do tobacco cessation counseling all know it is easier to get pregnant women to quit smoking, and that the earlier we talk to them the more effective it is. That may simply be because first trimester smoke smell and taste is difficult for most women. What we need is research on how to keep them as non smokers. We need to know for how long the different methods work.	We are developing journal manuscripts to aid in dissemination of the report findings. Unfortunately, there are few data on durability of interventions. We agree this is an important area of study and have noted it as a future research need.
Reviewer 9 (TEP)	General	This is an outstanding evidence report--the results are useful from a clinical and policy perspective, the target population and audience are explicitly defined, and the key questions appropriate and explicitly stated.	We appreciate your comments.
Reviewer 9 (TEP)	General	The report is an example of organization and clarity.	Thank you.
Reviewer 10 (TEP)	General	Comprehensive and well written	Thank you.
Reviewer 10 (TEP)	General	Page 11 - ES-3 line 47: "...minimum of 20..." It would be good to insert randomized after 20 to show that the sampling was random.	We also included prospective cohort studies for Key Questions 3, 4 and 5 so not all sampling was random.
Reviewer 11 (TEP)	General	The report is clinically relevant and should provide some new information to providers as they develop a treatment plan for the care of pregnant and postpartum women who smoke or have recently quit. This should also provide information for insurers, purchasers and health care systems. The term 'end user' should be reconsidered.	Thank you for the comments. The term end user reflects the broad range of readers who use these reports. Use of this term is not unusual for these reports.
Reviewer 11 (TEP)	General	Most providers will focus on the Executive Summary and Table 5 provides important evidence. The summary could be improved by providing additional tables or bullets under the intervention components in Table 5 that describe some general description the intervention. This could be modeled after the USPHS Treating Tobacco Use and Dependence guideline tables about treatment. This addition might be useful for clinicians in practice.	There is no room to add descriptions of the interventions to Table 5. We describe these in the text.

Commentator & Affiliation	Section	Comment	Response
Reviewer 12	General	Overall, this is an important study and should be used to instruct NIH and AHRQ to conduct further research in this area. There are too many poor and fair studies to put a great deal of reliance in the results. However, there are some areas that promise hope- incentives and mother-infant bonding- that should be researched further. It is surprising that breastfeeding was not controlled for in the analyses or addressed in the report. This is likely to interact with smoking in post-partum outcomes.	Thank you for your comment. The influence of breastfeeding on smoking outcomes would be a good area for future research.
Reviewer 12	General	The report does not address the safety of NRT or other smoking cessation drugs on the infant. It would be nice to see some reference to their safety.	We have added information in to the Introduction.
Reviewer 12	General	The HTML version is fraught with copy errors on several pages in which two words are typed next to each other with no space in between. Examples include Page 8, 2 lines down, mostpreventable; page 9 under KQ1: Postpartumfor; Page 10, 5 lines from bottom: Mostnon-English ; Page 12 , 16 lines up from bottom: Forconsistency;page 14 first line by KQ3: identifiedfour; Page 17 15 lines up from the bottom: severalinterventions; and on and on.. Please spell check the entire report for these errors.	These errors are not present in our version of the document. These are likely problems with conversion of the word document by the manuscript reviewer service.
Reviewer 12	General	The conclusions are not helpful in terms of future practice-based research that needs to be conducted. Policy cannot be made without a good research base. Clinical practice decisions cannot be made without a good research base. There are many promising new studies in this report (usually only one of each) that should be replicated, especially among low income women.	We have addressed the need for future research, including replication of promising studies, in the Discussion section.
Reviewer 13	General	Why focus only on RCTs? I didn't see an explanation.	RCTs represent the strongest evidence. However, they have limitations thus we included other study types for Key Questions 3, 4, and 5.
Reviewer 13	General	Is there a definition for "clinically meaningful differences"? need for a glossary of terms?	Clinically meaningful differences are those of sufficient magnitude to have practical relevance. We recognize this is somewhat subjective. We did not note recurring comments about unclear terms warranting a glossary.

Commentator & Affiliation	Section	Comment	Response
Reviewer 13	General	I was happy to see in the first summary that there is apparent effectiveness; however, I got a different sense from each of the subsequent summaries. Issue here is that the summaries seems to differ from each other; shouldn't they be consistent?	We have reviewed the summaries of the strength of evidence and findings in the Executive Summary for consistency.
Reviewer 13	General	As you know, this was requested because of the CHIPRA (Medicaid/CHIP) quality measures issues. Knowing this, it would have been helpful to know more about the characteristics of the interventions, providers, settings, and populations, specifically the applicability of the studies to the Medicaid/CHIP populations. E.g., in table 26.	We have described these characteristics as much as possible. Studies frequently lacked detail, especially in regard to interventions and providers.
Reviewer 13	General	Overall, a really excellent piece of work and I look forward to sharing the final version widely.	Thank you.