Comparative Effectiveness Review Disposition of Comments Report

Research Review Title: *Management of Infertility*

Draft review available for public comment from April 3, 2018, to May 1, 2018.

Research Review Citation: Myers ER, Eaton JL, McElligott KA, Moorman PG, Chatterjee R, Zakama AK, Goldstein K, Strauss J, Coeytaux RR, Goode A, Borre E, Swamy GK, McBroom AJ, Lallinger K, Schmidt R, Davis JK, Hasselblad V, Sanders GD. Management of Infertility. Comparative Effectiveness Review No. 217. (Prepared by the Duke Evidence-based Practice Center under Contract No. 290-2015-00004-I.) AHRQ Publication No. 19-EHC014-EF. Rockville, MD: Agency for Healthcare Research and Quality; May 2019. Posted final reports are located on the Effective Health Care Program search page. DOI: https://doi.org/10.23970/AHRQEPCCER217.

Response to Peer and Public Comments on this Research Review

The Effective Health Care (EHC) Program encourages the public to participate in the development of its research projects. Each research review is posted to the EHC Program Web site or AHRQ Web site in draft form for public comment for a 3-4-week period. Comments can be submitted via the Web site, mail or email. At the conclusion of the public comment period, authors use the commentators' submissions and comments to revise the draft research review.

Comments on draft reviews and the authors' responses to the comments are posted for public viewing on the 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.

#	Commentator & Affiliation	Section	Comment	Response
1.	Peer Reviewer #1	Quality of Report	Good	Thank you. No response needed.
2.	Peer Reviewer #4	Quality of Report	Good	Thank you. No response needed.
3.	Peer Reviewer #5	Quality of Report	Good	Thank you. No response needed.
4.	Peer Reviewer #6	Quality of Report	Good	Thank you. No response needed.
5.	TEP Reviewer #1	Quality of Report	Good	Thank you. No response needed.
6.	TEP Reviewer #2	Quality of Report	Good	Thank you. No response needed.
7.	TEP Reviewer #3	Quality of Report	Good	Thank you. No response needed.
8.	TEP Reviewer #4	Quality of Report	Good	Thank you. No response needed.
9.	Peer Reviewer #1	General	The 2nd sentence of the 2nd bullet point in "Key Messages" is confusing to me. I still don't know what is really meant by that sentence.	We have modified this sentence to clarify and to reflect the updated findings in the final report
10.	Peer Reviewer #1	General	--p12: The definition of infertility at 12 mos is outdated and should be updated as 12 mos for those <35 yo and 6 mos for those >35 yo.	We have made the suggested update.
11.	Peer Reviewer #1	General	--p20, ll 52: Adjunct treatments? Worthwhile explaining which ones rather than having the reader pull the referenced articles?	We have clarified which adjunct treatments within the key points
12.	Peer Reviewer #1	General	-p25: Increased rate of autism with IVF (low SOE). I would think listing the absolute difference in incidence would be more helpful in order not to throw out an inflammatory outcome such as this (without placing it in context, i.e., absolute incidence increase).	We have now included the hazard ratio for this increase within the key point.

#	Commentator & Affiliation	Section	Comment	Response
13.	Peer Reviewer #4	General	The meaning of this report is diminished by the snapshot of the time period evaluated. The conclusions need to be put in perspective to other findings. In isolation the conclusion can be be misleading.	The final report is now updated with literature through 2018.
14.	Peer Reviewer #4	General	Thank you for allowing me to review this document. This represents a large effort and the participants should be commended. It is difficult to synthesize and write.	Thank you. No response needed.
15.	Peer Reviewer #4	General	Minimize acronyms if possible	We have minimized acronyms when possible.
16.	Peer Reviewer #4	General	Overall the most difficult aspect of this review is that is based only on the articles in the time span. Some of these concepts I am trying to bring out (IE ICIS works for male factor infertility) can be mentioned in the discussion, thus avoiding the effect that this snapshot of literature does not give the "complete" picture. Specific articles may not suggest standard or care. IE start with ICIS works for male factor in the introduction and then state looking at subgroups IE ICIS vs IMSI. The statements are missing the big picture.	The final report is updated to reflect more recent evidence. In addition the table in the Discussion of current guidelines from SRM and NICE is meant to help summarize current standards and put out findings in the broader context.
17.	Peer Reviewer #4	General	A statement suggesting the consequences of excluding a that a large number (562) articles because of lack or reporting by underlying diagnosis should be included. Is it possible that many of those manuscripts could have address general aspects of fertility treatment. Perhaps better to make a comment of treatment in general with caveat that cannot comment on sub etiology rather than exclude?	This was discussed with the TEP and is admittedly a trade-off. Given the desire to organize the treatment choices by diagnosis, lack of specific details on diagnosis-specific outcomes for a specific treatment is an inherent limitation, particularly since a measure of association averaged across all diagnoses may over- or underestimate the treatment effect for specific diagnoses. We discuss the limitation of this restriction and its impact on our findings and generalizability extensively in the main report on page 137.

#	Commentator & Affiliation	Section	Comment	Response
18.	Peer Reviewer #4	General	A discussion point should be considered that while not optimal, ongoing pregnancy rate is very correlated to live birth rate. While it is a surrogate outcome, it is very close (with the exception of late miscarriage). I agree a push should continue to report live birth, but good literature should not be removed fully.	This point was discussed with the TEP, who agreed that limiting to live birth was appropriate; we have expanded the discussion of the impact of this restriction .
19.	Peer Reviewer #4	General	A bit too much on theoretical issues like time trade off and CEA.	We have removed the discussion related to CEA and time tradeoff from the evidence summary but have retained this discussion within the more expansive main report.
20.	Peer Reviewer #4	General	I understand this is a draft but numbering of references and tables should be evaluated as some are not correct.	We have corrected the numbering of references and tables.
21.	Peer Reviewer #5	General	Review verb tense throughout paper	We have edited the final report and verified verb tense.
22.	Peer Reviewer #5	General	This may be my ignorance, but I am trying to grasp the objective of this paper. If the objective is for it to helpful for clinicians, it is quite dense and long, and may be beneficial to be broken into several papers. If used for policy, etc., it may be more appropriate	The goal of the systematic review is provide a complete review of the available evidence which may be helpful to the decision making of diverse stakeholders. The reviewer is correct that papers which focus on specific key questions or the impact on specific perspectives may be beneficial to individual stakeholder groups and will be produced as well.
23.	Peer Reviewer #5	General	What is the goal of the lengthy summary at the beginning of the paper? Again, this may be a formatting instruction for this type of document.	The evidence summary at the beginning of the report is designed to provide an overview of the full report. It does not provide the details of the underlying evidence provided in the main report but instead focuses on the key points, synthesis of the evidence, and putting the findings in to context.

#	Commentator & Affiliation	Section	Comment	Response
24.	Peer Reviewer #6	General	The report focuses solely on medical interventions and neglects the things a couple might do on their own to achieve pregnancy such as weight loss, cycle charting, timing of intercourse, and supplements.	The interventions were suggested by the TEP. In part because the review focused on management after a diagnosis had been made, some interventions (such as cycle charting, intercourse timing, counseling on coital frequency, etc.) that are frequently done prior to a formal diagnosis were not included. Interventions such as weight loss (particularly for PCOS) or dietary supplements were included if they otherwise met inclusion criteria. We have clarified this restriction in the overview of the scope of the report.
25.	Peer Reviewer #6	General	#1 reduces the generalizability of their report to almost all couples but particularly those who lack the resources for expensive and invasive treatments.	As described above, the interventions were suggested by the TEP. In part because the review focused on management after a diagnosis had been made, some interventions (such as cycle charting, intercourse timing, counseling on coital frequency, etc.) that are frequently done prior to a formal diagnosis were not included. Interventions such as weight loss (particularly for PCOS) or dietary supplements were included if they otherwise met inclusion criteria.

#	Commentator & Affiliation	Section	Comment	Response
26.	TEP Reviewer #1	General	In general, this is a fairly complete review of the literature which is addressing the right outcomes to the right audience. The key questions are appropriate and explicitly stated. The relative paucity of strong RCTs leads to rather weak strengths of evidence and therefore, the conclusions are also fairly weak. This limits the strength of clinical guidance that is suggested but that is not the fault of the authors.	Thank you. No response needed.
27.	TEP Reviewer #2	General	I don't understand why the letrozole versus clomiphene message was presented first as a key message. It's not clear to me that this was the most important finding from these analyses	We have revised the ordering of the key messages. The key message ordering is not necessarily indicative of clinical importance. In this case, the message was presented because it was one of the few with reasonable strength of evidence
28.	TEP Reviewer #2	General	I think the most important message is the last. We need better consistency in reporting to allow better treatment comparison	We have revised the order of the key messages as suggested.
29.	TEP Reviewer #3	General	Overall the report is a nice addition to the literature. It's well written, easy to follow and clinically meaningful. I thought the research recommendations were lacking although the rationale was logical. From a public health surveillance perspective, it might be nice if authors could comment on what additional data fields would be helpful in addressing gaps in the literature. Linking cycles across patients and moving toward cumulative success rates is under way. Are there other things we should be doing?	We have added additional specific suggestions (especially diagnosis-specific outcomes for studies which include multiple infertility causes). Even if an individual study is underpowered, reporting of these subgroup results can facilitate subsequent pooled analyses or meta-analyses.
30.	TEP Reviewer #4	General	As a practicing Reproductive Endocrinologist, this report is not surprisingly discouraging - not in the research strategy or implementation (superior in my opinion) but rather in the limited guidance (12 points) that the document will provide to those managing the care of infertility patients.	Thank you for your comments. We agree that there is a large amount of uncertainty remaining in the optimal treatment of infertility patients.

#	Commentator & Affiliation	Section	Comment	Response
31.	TEP Reviewer #4	General	As stated in the Conclusion section of the Structured Abstract, there are only two supporting strategies (PCOS and unexplained infertility) that fall out of this report. Given the limited advice that the report can convey based on the limited strength of evidence of the studies, I believe that the authors should attempt to manage the reading audience expectations from the very beginning. For example, use language found on page ES18 and in the first few pages of the Introduction in the Structured Abstract, reminding the audience about the complexity of the problem(s) doing research in this arena. There are two patients involved, there are non-homogeneous conditions (obese PCOS, vs thin PCOS, all different stages of endometriosis, etc), multiple outcomes (pregnancy vs live birth vs ectopic) etc,.	We are somewhat constrained in length and content for the Structured Abstract, but have revised to reflect the reviewer's suggestion.
32.	TEP Reviewer #4	General	I believe that the KQ's were adequately defined and provide an opportunity to "set the stage" for disappointment - look at KQ1a, KQ2a etc - most readers already know the answer is "yes" to almost every variable listed. There is so much confounding going on that clear answers will be few and far between.	Thank you. No response needed.
33.	TEP Reviewer #4	General	Lastly, I had to look up "IMSI" because the term had no clinical relevance to me. This procedure is not currently used in clinical IVF and as such, the finding that it isn't beneficial is probably self evident.	The interventions considered as part of the "PICOTS" were inclusive of TEP input. We have added context about the use, or lack thereof, of IMSI in current US practice.
34.	Jessica Goldstein (Public Reviewer #1)	General	Taking on the topic of infertility, a single systematic review is akin to attempting to take on a topic such as cancer in a single document. There are many sub-topics under the umbrella term and any one of which could command its own systematic review. The risk of aggregating all infertility topics in a single monograph is diluting the overall value by having great length but insufficient depth.	We agree that this report is very broad and expansive in its goal of evaluating the evidence across infertility diagnoses. We imagine that several readers may only focus on the findings of a specific key question which is pertinent to their population of interest.

#	Commentator & Affiliation	Section	Comment	Response
35.	Jessica Goldstein (Public Reviewer #1)	General	The key messages in the review are general and there are few concrete recommendations. Left without clear direction on a number of topics and the absence of comments on others practitioners may not implement meaningful changes in practice thus muting the impact of the review.	The purpose of the AHRQ report is to review and report the available evidence related to the management of infertility. The translation of these findings in to clinical recommendations is a separate goals for policymakers.
36.	Jessica Goldstein (Public Reviewer #1)	General	Though the review is lengthy and well referenced, some conclusions drawn on the basis of low SOE do not seem plausible.	For the final report we have reviewed all SOE ratings and confirmed the link between these ratings and any conclusions drawn.
37.	Jessica Goldstein (Public Reviewer #1)	General	The goal of this document and all documents produced by the EPCs is to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. The draft in its current form may not achieve the target goal.	We thank the reviewer for their comments and hope that the final report will aid decisionmakers in their review of the available evidence.
38.	Jessica Goldstein (Public Reviewer #1)	General	The document is too dense/long to glean meaningful information in a concise manner. The document is helpful in looking at each reference per key question but the typical reader will not look through 150+ pages of this but may look to get a sense of literature to date.	We understand that the document is long. The executive summary provides a shorter summary of the main findings which may be helpful to specific readers who are not interested in the full evidence review.
39.	Jessica Goldstein (Public Reviewer #1)	General	This report would certainly not represent current medical thinking or relevant data in this subject area.	Thank you for your response. We have addressed specific concerns individually.
40.	Jessica Goldstein (Public Reviewer #1)	General	I am puzzled as to how studies were chosen, what criteria was used, as well as the interpretation of the data?	The inclusion and exclusion criteria as well as the search strategy and screening steps are detailed in the methods section of the report.

#	Commentator & Affiliation	Section	Comment	Response
41.	Anonymous (Public Reviewer #2)	General	I like the key findings but don't know about the SOE. - feel like this report is more clinician focused not patients focused. Is it possible to make evidence summary document more patient focused (without SOE)	The key messages at the beginning of the report are presented without the SOE ratings but we have retained the SOE ratings within the key points in the main report.
42.	Anonymous (Public Reviewer #2)	General	Table numbering is wrong (table number mismatch in title and in the text)	We have corrected table numbering and mismatches.
43.	Peer Reviewer #1	Introduction	Adequate as is.	Thank you. No response needed.
44.	Peer Reviewer #2	Introduction	This manuscript is intended to provide a comprehensive systematic review of the safety and effectiveness of infertility treatments for a variety of patient populations. The analytic framework is helpful in providing an overview of the questions, study populations, outcomes of interest and confounders/modifiers. The authors have identified the principle modifiers including age, ovarian reserve and BMI in this framework. The intent is also to capture efficacy but also to assess safety/adverse effects including ohss, cancer, effects on offspring, etc. This obviously represents a huge amount of work and the authors follow strict criteria for conducting systematic reviews.	Thank you. No response needed.
45.	Peer Reviewer #4	Introduction	The key questions are well laid out. However, the sub questions are not really addressed (i.e. treatment by age and other cofactors.) It may be best to say the desire was to look at outcomes in this patient characteristics, but due there was a paucity of data to evaluate those questions. Report when there was information, rather than reporting when there was no information specific to those groups.	We now describe in the description of included studies how the findings about these patient subgroups was minimal but now we highlighted in the report those cases where findings in these specific subgroups was possible.

#	Commentator & Affiliation	Section	Comment	Response
46.	Peer Reviewer #5	Introduction	I do not believe that advanced reproductive age is the most common cause of infertility, as ovulatory dysfunction and male infertility are more common. "Age related fertility decline" may be a more appropriate term, and authors may review how common it is among patient presenting for treatment. (Page 12, line 24)	We have clarified that this is the most common demographic factor, not the most common single cause.
47.	Peer Reviewer #5	Introduction	It seems like the authors are using PCOS and ovulatory dysfunction synonymously. Although PCOS may be common, by just using PCOS, you are not including those with other issues, such as hypothalamic amenorrhea, hyperprolactinemia. The authors may choose to only study PCOS, however, considering your other key questions, it seems like every other cause if infertility is being studied, except other forms of ovulatory dysfunction. The authors should include other causes or mention why they are not in the text. (Page 13)	The decision to limit to PCOS rather than other ovulatory disorders was made in consultation with the TEP, based primarily on frequency of diagnosis. We have added other causes of infertility, as well as a discussion of the rationale for excluding them.
48.	Peer Reviewer #5	Introduction	Otherwise, background is a good summary of several of the issues and challenges associated with infertility treatment	Thank you. No response needed.
49.	Peer Reviewer #5	Introduction	Consider adding some risks of LOD (surgical risks this this population, adnexal scarring, ovarian reserve…)	We have added the general and specific surgical risks associated with LOD
50.	Peer Reviewer #6	Introduction	They completely gloss over the concept of 'subfertility' and instead focus on the medical definition of 'infertility' with its dichotomous limitations. The authors freely acknowledge the limitations of that approach when they cite how many people will conceive after one year of subfertility.	We have expanded on this description; again, the approach specified by the TEP was focused on post-diagnosis management in the US. We agree that subfertility is a valuable concept—it may be even more appropriate to consider it not a diagnosis at all, but rather "normal" couples who happen to be in lower percentiles of the functional distribution of time to pregnancy.

#	Commentator & Affiliation	Section	Comment	Response
51.	TEP Reviewer #1	Introduction	ES-1- lines 16-23- Is this for the United states only? I am sure that utilization of infertility services vary widely around the world and even in the US depending on insurance coverage of care. Some mention of different utilization of services is warranted.	We have clarified that these statistics reflect the United States population.
52.	TEP Reviewer #1	Introduction	ES-4- analytic framework- a potential modifier is frequency of intercourse/ presence of sexual dysfunction which is not often mentioned but is a real issue clinically.	Although we agree that this is an important potential monitor, the TEP did not feel it was important to include. In part because the review focused on management after a diagnosis had been made, some interventions (such as cycle charting, intercourse timing, counseling on coital frequency, etc.) that are frequently done prior to a formal diagnosis were not included.
53.	TEP Reviewer #1	Introduction	p 22- In this table I see only one study comparing cc with metformin alone and this study (ref 124) concludes that clomid is superior to metformin alone. Why does a fair quality meta analysis trump a well done PRT? I think the conclusion of a higher pregnancy rate with CC vs metformin is pretty widely held belief based on clinical experience and the data.	The meta analysis included the individual study (ref 124 from the draft report) and combined with other studies does not a difference in live births. The low SOE for this finding has been retained.
54.	TEP Reviewer #1	Introduction	p25- I have trouble with lumping all studies of "LOD" together when there is not a good definition of how to perform LOD. There is wide variation in how this procedure is performed.	The issue of how to compare results of studies which evaluate a procedure that may be performed in a number of ways (and where provider skill may vary) is a recurring one without a satisfying resolution. We have added some discussion of this issue.
55.	TEP Reviewer #2	Introduction	Inclusion/exclusion are clear.	Thank you. No response needed.

#	Commentator & Affiliation	Section	Comment	Response
56.	TEP Reviewer #2	Introduction	Definitions are ok except for strength of evidence. That should be uniform for all single study conclusions.	We rated the strength of evidence using the approach described in the AHRQ EPC's Methods Guide which incorporates information on study quality, consistency, directness, precision, and reporting bias. In the final report, we confirmed consistency in terms of rating outcomes where evidence is supported by only one study.
57.	TEP Reviewer #2	Introduction	I think that if there is concern that there is bias, then the grade should be insufficient rather than putting in a result that may or may not be useful to providers.	Although potential risk of bias reduced the SOE, there are instances where available evidence from studies of varying quality and risk of bias could be rated as having an low or moderate SOE.
58.	TEP Reviewer #3	Introduction	Regarding the prevalence of infertility reported on page 12, lines 15-16, would it be appropriate to cite the more current information (2011-2015) from the NSFG website (6.7%)? https://urldefense.proofpoint.com/v2/url?u=https-3A__www.cdc.gov_nchs_nsfg_key-5Fstatistics_i.htm-23infertility&d=DwICaQ&c=imBPVzF25OnBgGmVOlcsiEgHoG1i6YHLR0Sj_gZ4adc&r=5RZkmyYg30xi-DHJP9Og3PUb_b9WGnyH29XDJS51Vwo&m=MxY3yLMJRHvwta3s1Gdypn1hjnvWwZBebwxOT2wCTjs&s=JZdoKw2SsfkwBHR4UozuM766pW_m9k47ESwm6NoQSYQ&e=	We have updated.
59.	TEP Reviewer #3	Introduction	On line 37 - consider explaining first that IVF can be done either with ICSI or without ICSI (conventional IVF). As written, it sounds like IVF and ICSI are two different things. Also the terminology used changes throughout the document. For example, on page 99, lines 47-48, the phrase "compared outcomes between IVF and combined IVF and ICSI" is used.	We have clarified the description.

#	Commentator & Affiliation	Section	Comment	Response
60.	TEP Reviewer #4	Introduction	I love the Introduction and as I recommended above, some abbreviated form of these "complexities" might be pulled into the Summary Abstract (which might be the most read portion of the manuscript).	We have pulled in to the structured abstract discussion of the complexities of research studies in this area.
61.	TEP Reviewer #4	Introduction	Figure 1 Analytic framework is amazing	Thank you. No response needed.
62.	Peer Reviewer #1	Methods	Sufficient	Thank you. No response needed.
63.	Peer Reviewer #2	Methods	The methodology is appropriate.	Thank you. No response needed.
64.	Peer Reviewer #4	Methods	Need a definition of strength of the evidence in the summary	We have added in the definitions of the SOE ratings to the evidence summary as requested.
65.	Peer Reviewer #5	Methods	The methodology seems sound, however the editors may consider evaluation by someone with more expertise in study design/statistics	Thank you. No response needed.
66.	Peer Reviewer #5	Methods	Appreciate the outcome of live birth! I agree that this is the most important with infertility treatment studies	Thank you. No response needed.
67.	Peer Reviewer #6	Methods	Methodology was stellar.	Thank you. No response needed.
68.	TEP Reviewer #1	Methods	I think this is good except that some very rare outcomes including neonatal death and birth defects, the studies included had no chance to detect any differences, so it makes little sense to include them in the tables, in my opinion.	We agree with the limits of existing studies to detect some outcomes of interest, however these outcomes were identified by our stakeholders and technical expert panel as highest priority and therefore the data (however limited) is presented in our synthesis.
69.	TEP Reviewer #2	Methods	Reader needs to understand why 95% confidence intervals are important. Would be nice if there was a sentence about the analysis methodology in this (ANOVA? ANCOVA?)	We have inserted in to the methods some clarification as to the interpretation of 95% confidence intervals.

#	Commentator & Affiliation	Section	Comment	Response
70.	TEP Reviewer #2	Methods	Summary of studies should have a little more information. Studies were from US? Canada? Multicenter?	Although this information is not provided in the Evidence Summary, it is provided in the main report for each key question.
71.	TEP Reviewer #2	Methods	Studies that are not multicenter should be noted in this report as these studies are have a higher risk of bias than multicenter studies.	The individual components of our quality and risk of bias ratings are included in the Appendix.
72.	TEP Reviewer #3	Methods	I imagine this has already been addressed but it seems odd that the review only goes through September of 2015 when we are currently in 2018. Are there plans to update the search prior to publication?	The final report includes evidence through 2018.
73.	TEP Reviewer #4	Methods	Clearly defined, logical inclusion/exclusion criteria. The search strategies were clearly explained. Outcome measures chosen are appropriate. Strength of evidence section was very readable and useful. No real statistical applications in this paper.	Thank you. No response needed.

#	Commentator & Affiliation	Section	Comment	Response
74.	Anonymous (Public Reviewer #2)	Methods	Page 7: please be clear about your inclusion of SRs - it is kind of confusing	We have clarified in our methods and throughout the report the role of existing systematic reviews in the findings and our SOE ratings. While systematic reviews and meta-analyses were not study designs qualifying for inclusion and abstraction under our screening criteria, we did flag relevant articles of these types as part of the screening process. Component references from these systematic reviews were reviewed and when studies met our inclusion criteria, they were included in our report. For systematic reviews which were identified as relevant to the individual KQs but included mostly studies prior to 2007, we summarize the findings from these existing reviews and the consistency of their findings with those from our included studies in the appropriate results sections.
75.	Peer Reviewer #1	Results	There is a recent abstract presented at ASRM (I will attach this) by the RMN that could really change the landscape of superovulation for unexplained infertility and it might be worthwhile to wait to see if this is published (I would almost guarantee it would be).	We reviewed this abstract during the update for our final report but since the findings have not yet been published as a full manuscript – they were excluded from our review.

#	Commentator & Affiliation	Section	Comment	Response
76.	Peer Reviewer #2	Results	As the authors acknowledge, the challenge with the literature in infertility is that often data are lacking or insufficient to adequate address the questions of interest. Thus, the final summary tables presented in the beginning of the document appear to be somewhat haphazard and do not seem to capture all of the pertinent clinical questions and leaves the reader somewhat perplexed. Moreover, there are no clear summary statements to inform clinical care (what is the take home point?). For example, it would be helpful to conclude the section on PCOS and infertility with a statement: “Letrozole is recommended as a first line therapy for ovulation induction given evidence of efficacy and safety”. I recognize that more detail is presented farther in the document, but the summary statements are not phrased in a way that directs clinical care. Compared to the ASRM and NICE guidelines, these recommendations are not as directive.	The purpose of the AHRQ systematic review reports is to summarize the available evidence and the strength of evidence for specific outcomes of interest across comparative interventions. The systematic reviews do not intend to provide guideline directives but instead present evidence which might be informative to guideline groups.

#	Commentator & Affiliation	Section	Comment	Response
77.	Peer Reviewer #2	Results	The sections on endometriosis and tubal/peritoneal faction are extremely limited and only address 1 question – single vs. double embryo transfer because the SOE was limited in all other areas. I would argue that it is limited even for this question since potential HARM is not addressed. As a clinician, I would read this document and change my practice to double ET for all patients with endometriosis or tubal factor. However, what is the risk of multiple pregnancy? Again, a summary statement with more nuanced information and recommendations for practice would be extremely helpful UP FRONT. Alternatively, leave these analyses out in the beginning of the document and make more directive suggestions.	The findings are limited by the studies which met inclusion criteria, and also by the varying strength of evidence for specific outcomes. The purpose of the review is not to provide specific recommendations for practice, but rather to summarize the available evidence so that specialty societies, payers, and other policy makers can make specific practice recommendations. Throughout the report we highlight the evidence supporting any benefits and harms of our included outcomes of interest. When not available, we note that the evidence base does not include evidence for those specific outcomes of interest.
78.	Peer Reviewer #2	Results	The unexplained infertility section is very limited by the fact that age does not seem to be taken into account. There are data suggesting that IVF should be the first line therapy for unexplained in fertility in women 38 years of age and over. Also, duration of infertility is an important consideration in treatment decisions and has been shown to modify outcomes. The summary tables are far too simplistic to be of any real clinical utility. Again, how should these data be interpreted when applying them to patient care?	We have included the evidence from the FASTT trial which provides evidence on the impact of firstline IVF.

#	Commentator & Affiliation	Section	Comment	Response
79.	Peer Reviewer #3	Results	The section on male infertility is surprisingly short and does not address any actual treatment of male factor infertility. It only addresses the use of ART for male infertility. This was not what I had envisioned with this project was started (looking back at the original document, it seemed that treatment of male infertility as a key question would have involved more than just ART). I have also found that many of the studies are small, and often of questionable quality. Finally, the Practice Guidelines Committee of the American Society for Reproductive Medicine has produced multiple documents that cover all of the key questions, including the evaluation and treatment of male infertility. These guidelines, which are typically evidence-based, are much cleaner, and make more definitive statements. I am concerned that this document is much too long without clear statements, thus making it less useful for the practitioner.	The systematic review has been updated to include additional studies published more recently. Unfortunately included studies did not evaluate other treatments. Note that for the majority of screened abstracts the outcomes assessed related to semen parameters rather than live birth or other outcomes of interest to this report. We do include a summary of the ASRM guidelines in the discussion. Note that the systematic reviews do not intent to provide guideline directives but instead present evidence which might be informative to guideline groups.
80.	Peer Reviewer #3	Results	While the question states safety and effectiveness of available treatment strategies, the report does not discuss any treatment of male factor other than ART.	Unfortunately included studies did not evaluate other treatments. Note that for the majority of screened abstracts the outcomes assessed related to semen parameters rather than live birth or other outcomes of interest to this report.
81.	Peer Reviewer #3	Results	I'm not certain how the key points were determined, but the evidence for Vit E or Zn supplementation does not really make sense from a clinical perspective. It is interesting that the key points listed do not really say much about the treatment of actual male factor infertility.	We agree that the evidence supporting male factor infertility treatments is sparse. We are limited to studies which reflect our inclusion and exclusion criteria. We do summarize findings from ASRM in the discussion and highlight difference from our systematic review of the evidence.

#	Commentator & Affiliation	Section	Comment	Response
82.	Peer Reviewer #3	Results	It appears that the literature searches regarding ART for male factor missed some important studies. There are several randomized trials of varicocele repair vs IVF, as well as numerous meta-analyses. There are not any articles cited regarding the treatment of azoospermia, either obstructive or non-obstructive. Similarly, there are multiple studies demonstrating no significant difference in outcome for IUI utilizing gonadotrophins vs clomid/letrozole. Perhaps these studies did not fit the search criteria used.	The final report has been updated to include literature through 2018 which now includes 26 studies related to male infertility.
83.	Peer Reviewer #3	Results	I am truly puzzled as to how studies were chosen, what criteria was used, as well as the interpretation of the data. It appears that according to this report, ART, including ICSI, does not seem to have good evidence in the treatment of male factor infertility. Non-specific treatments, such as anti-oxidant use, while studied, are not specific enough to be able to say that their use makes a significant difference.	The methods related to inclusion and exclusion criteria and our relevant interventions and outcomes of interest are detailed in the methods section of the report.
84.	Peer Reviewer #4	Results	I suggest that OR or RR should be placed in the tables when possible. They are listed sometimes and not others (IE in summary table A, they are listed for some outcomes, but not others.) In the main body they are not listed at all, and only P values are listed. This can be misleading as data can suggest a finding quite convincingly while the P value is NS. For example, ovarian drilling compared to medical intervention. The conclusion that they are not different is “correct” based on the CI, but the consistent direction of the RR and the P values close to SS suggests otherwise. A reader should be able to make this interpretation based on data in the table. This is especially true as meta-analysis was not performed.	We have modified the tables to be more consistent in terms of whether we report summary estimates and p values.

#	Commentator & Affiliation	Section	Comment	Response
85.	Peer Reviewer #4	Results	The conclusion for endometriosis is misleading. The fact that two embryos gives a higher pregnancy rate than on embryo is likely true for all causes of infertility, not just endometriosis. The conclusion as written suggests that the only treatment women with endometriosis is to transfer two embryos. (not that many treatments can work, or IVF has success rates comparable to other causes of infertility). That will lead to bad medicine given the trend to decrease the number of embryos in all ART. There are papers looking at outcome of IVF with different dx and if women with endometriosis have a higher or lower pregnancy with IVF. Perhaps these studies are before and after your dates for review.	We have clarified that this finding is true across all causes of IVF, adding to the key points for each KQ. "As with other indications for IVF, use of single-embryo transfer is associated with slightly lower live birth rates but significantly reduced multiple gestation rates (low SOE)."
86.	Peer Reviewer #4	Results	The same can be said for conclusion regarding treatment of tubal factor infertility. The first finding or statement should be that IVF works for this group of women. The finding that two embryos is better than one is misleading and suggesting that more aggressive therapy is needed in this group.	We have clarified that this finding is true across all causes of IVF, adding to the key points for each KQ. "As with other indications for IVF, use of single-embryo transfer is associated with slightly lower live birth rates but significantly reduced multiple gestation rates (low SOE)."
87.	Peer Reviewer #4	Results	These examples are not tempered by the next statement the multiple embryo transfer can result in higher multiple birth as that outcome is not pre-specified.	We have clarified that this finding is true across all causes of IVF, adding to the key points for each KQ. "As with other indications for IVF, use of single-embryo transfer is associated with slightly lower live birth rates but significantly reduced multiple gestation rates (low SOE)."

#	Commentator & Affiliation	Section	Comment	Response
88.	Peer Reviewer #4	Results	The statement of the balance of number of embryos listed in finding applicable across all key question should be highlighted and maybe come before the specific recommendations for each Q.	We have clarified that this finding is true across all causes of IVF, adding to the key points for each KQ. "As with other indications for IVF, use of single-embryo transfer is associated with slightly lower live birth rates but significantly reduced multiple gestation rates (low SOE)."
89.	Peer Reviewer #5	Results	May want to mention difference in letrozole/clomid at different BMI (page 16). This is in reference to citation 44. This is mentioned later in the paper.	We have added a statement regarding BMI regarding the Legro study as suggested.
90.	Peer Reviewer #5	Results	Should add the age concern with reference 94—(I see this mentioned on page 27…)	We discuss the relevance of the reference, given the dates of the included studies, with the fuller discussion of the review.
91.	Peer Reviewer #5	Results	Seems like a strong statement to say IVF associated with autism and ovarian malignancy (page 25).	Both of these findings have low SOE and therefore uncertainty remains.
92.	Peer Reviewer #5	Results	I would not say there is a "slightly" lower chance with an SET—it may be about 10% lower (page 25)	We have modified to say simply "lower."
93.	Peer Reviewer #5	Results	It is not surprising that all patients with a 2 embryo transfer have a greater chance of live birth. Please be sure to mention the associated increased rate of multiple gestation.	Our findings do emphasize the increased rate of multiple gestation with 2-embryo transfer.
94.	Peer Reviewer #5	Results	The results section in the body of the paper (not the summary) is very comprehensive and I appreciate that they mention special populations. This is critical due the heterogeneity of the study populations.	Thank you. No response needed.
95.	Peer Reviewer #5	Results	Page 98—in regards to the FASST trial—were the delivery costs different due to the incidence of multiple gestation? If so, it may be worth mentioning.	We now clarify that multiple birth rates did not differ significantly between the two arms.
96.	Peer Reviewer #5	Results	May want to mention if SART is doing more prospective data gathering on donors (page 30)	We were unable to locate any information on formal efforts to track long-term outcomes of donors.

#	Commentator & Affiliation	Section	Comment	Response
97.	Peer Reviewer #6	Results	The tables are quite helpful and clear. I know of no ommissions.	Thank you. No response needed.
98.	TEP Reviewer #1	Results	p 22- In this table I see only one study comparing cc with metformin alone and this study (ref 124) concludes that clomid is superior to metformin alone. Why does a fair quality meta analysis trump a well done PRT? I think the conclusion of a higher pregnancy rate with CC vs metformin is pretty widely held belief based on clinical experience and the data.	The meta analysis included the individual study (ref 124 from the draft report) and combined with other studies does not a difference in live births. The low SOE for this finding has been retained.
99.	TEP Reviewer #1	Results	p25- I have trouble with lumping all studies of "LOD" together when there is not a good definition of how to perform LOD. There is wide variation in how this procedure is performed. Were all studies included using the same procedure- ie same energy source, same technique and number of areas in the ovary "burned" etc. This lack of precision should be mentioned as an issue.	We have highlighted in the text how the studies differed in terms of the surgical methods used. When possible the quantitative synthesis was comparing surgical management (as a class) compared with oral agents.
100.	TEP Reviewer #1	Results	p.51- line 15-16- what is uterine perturbation? Is this "endometrial scratching"? I think most will not know what this paper is referring to without more of a description.	We have clarified the description of the specific intervention from the included study
101.	TEP Reviewer #1	Results	p.52- line 40-44. This sentence is not true. couples did go to IVf without first getting gonadotropins in one arm of the study.	Thank for noting the error. We have clarified
102.	TEP Reviewer #1	Results	In general, I have a comment about the even mentioning outcomes of neonatal deaths or congenital anomalies when the subject of much of this review are relatively small PRTs. these rare outcomes are mentioned in the table on page 54 for example for outcomes of small PRTs and it does not really add anything to this paper.	Given than these outcomes were pre-specified they are reported within the report regardless of how sparse the evidence was. Our rating of the strength of evidence and inclusion of the findings in the key points/messages etc. are focused on the outcomes with sufficient evidence/findings.

#	Commentator & Affiliation	Section	Comment	Response
103.	TEP Reviewer #1	Results	p63- I see reference 120 used often in looking at IVF outcomes for different diagnostic groups. I think this could be problematic since reporting of diagnostic groups in the SART database is notoriously poor and the definitions as to how the diagnosis was made is poorly defined. this should be mentioned as a weakness.	We have added a sentence noting limitations related to diagnosis in the SART database
104.	TEP Reviewer #1	Results	p 63 table 27 lines 47-52- under conclusion- what are you trying to say here? was there no difference or was the pregnancy rate higher after double embryo transfer? in this sentence you say both	This row has been clarified and also updated to reflect additional evidence from the final report.
105.	TEP Reviewer #1	Results	p 67 line 42-53- the use of IMSI (a relatively rare procedure of no consequence) is not reported in the SART database so this sentence has to be wrong.	We have corrected this error.
106.	TEP Reviewer #1	Results	Why are you reporting on IMSI anyway? If we are going to get into specifics of IVF, there are many technical aspects of this that have been studied in multiple studies- assisted hatching, ICSI in the absence of male factor, extended culture of embryos, PGS etc. I do not think you want to get to this level of detail in the IVF section unless you are going to delve into many aspects of this procedure.	We have clarified in the report the clinical use of IMSI but the reviewer is correct that we do not go in to the specific detail of the various IVF techniques.
107.	TEP Reviewer #1	Results	p.74- line 28- It seems to me that your statements regarding increases in ovarian cancer are not consistent. Can you please review? I think it would be helpful to define these are true invasive cancers versus LMP tumors since the prognosis is so different. In some sections, the risks do not seem to be higher when controlling for nulliparity, duration of infertility etc, and then in other areas you state it is higher.	We have revised the description of ovarian cancer findings, which, as the reviewer notes, are complex.
108.	TEP Reviewer #1	Results	p 74- line 30-31- is there really only one study of autism that meets your inclusion criteria-? please review. The literature as a whole is very mixed on this topic to my review.	Unfortunately the evidence on this topic is sparse and only one study continued to meet our inclusion criteria.

#	Commentator & Affiliation	Section	Comment	Response
109.	TEP Reviewer #1	Results	p. 74 line 35-37- the evidence that SET reduces multiple birth has strong SOE. There are multiple studies demonstrating a dramatic reduction in multiple birth- many more studies than you have cited.	We have revised the report to reflect the updates in the literature but unfortunately still identified sparse evidence meeting our inclusion and exclusion criteria.
110.	TEP Reviewer #2	Results	Table and figures were clear. No studies were overlooked that I am aware of. See my previous comment about single center trials. Also, any studies that did not report their randomization should be considered suspect.	Thank you. No response needed.
111.	TEP Reviewer #2	Results	Also, I think the results should be reported as to whether the publications were US or ex US. Ex US publications may not provide similar results because practices of ART may be significantly different	The geographic location of each study is included in Appendix E.
112.	TEP Reviewer #3	Results	Page 59, Table 9. CDC's data collection system is called NASS - National Assisted Reproductive Technology Surveillance System. (not Artificial Reproductive Technology)	We have corrected the typographical error
113.	TEP Reviewer #4	Results	Figure 2 is clear (depressingly clear that more research is needed)	Thank you. No response needed.
114.	TEP Reviewer #4	Results	page 20 line 48 - should "selective estrogen receptor modulator" be used instead of "anti- estrogen" for clomiphene citrate, perhaps throughout the paper?	We have made the suggested change throughout the report
115.	TEP Reviewer #4	Results	Page 21 line 5 - "Table 4" should read "Table 10" I think that many (most) of the Table numbers are mislabeled	We have corrected the table numbering throughout.
116.	TEP Reviewer #4	Results	I think that the Figures and Tables are easy to read and self explanatory	Thank you. No response needed.
117.	Jessica Goldstein (Public Reviewer #1)	Results	The Endometriosis section: key finding that 2-embryo transfer had higher live-birth rate (LBR) per cycle seems to support double-embryo transfer in this population while ignoring the risk and negative impact of twin gestation.	Multiple gestation is an outcome of interest however the evidence did not support findings regarding this outcome. We have however highlighted this risk in our discussion of the embryo transfer evidence.

#	Commentator & Affiliation	Section	Comment	Response
118.	Jessica Goldstein (Public Reviewer #1)	Results	Tubal factor: authors again appear to endorse two-embryo transfer as strategy to improve LBR	We have clarified that there is a trade-off between live birth rate and reduction in multiple births for this KQ.
119.	Jessica Goldstein (Public Reviewer #1)	Results	Elective single-embryo transfer (eSET) discussed in key questions re: lower LBR but significant reduction in multiple rate (low SOE) - why wasn't this mentioned in previous section endo/tubal in terms of single versus double ET?	We have clarified that there is a trade-off between live birth rate and reduction in multiple births for this KQ.
120.	Jessica Goldstein (Public Reviewer #1)	Results	Male factor: overall. appears most of this evidence was low SOE with exception of ICSI vs IMSI, which is less clinically relevant.	We agree that the evidence supporting male factor infertility treatments is sparse. This continues to be the case in the final updated report.
121.	Jessica Goldstein (Public Reviewer #1)	Results	While the abstract suggests that they are investigating treatments for "couples with male factor infertility," there are almost no data evaluated for male infertility. Indeed, there was no comparison of specific male-factor treatment compared to assisted reproductive technology (ART).	The updated report include additional studies related to male factor infertility, however the data does continue to be sparse and be an area of needed future research.
122.	Jessica Goldstein (Public Reviewer #1)	Results	The management of RCTs for antioxidant therapy for male factor is superficial and indirect (as well as contradictory), and there is no mention of hormonal therapy, varicocele treatment, or management of obstructive azoospermia beyond sperm retrieval and in vitro fertilization (IVF).	Studies included reflect the interventions, comparators, and outcomes of interest prioritized through our key informant and TEP processes.
123.	Jessica Goldstein (Public Reviewer #1)	Results	All Key Questions offer a more distilled and helpful take-away, but it is unclear why the ICSI-autism link was included if low SOE.	We include in the key points those findings where SOE was rated as low, moderate, or high.

#	Commentator & Affiliation	Section	Comment	Response
124.	Jessica Goldstein (Public Reviewer #1)	Results	A large concern is that maternal age was barely mentioned.	Age was a subpopulation of interest and we looked for evidence for such populations throughout the report. We report findings for this subgroup in the unexplained infertility diagnosis section which is the only diagnosis where this subgroup was specifically evaluated for outcomes of interest.
125.	Jessica Goldstein (Public Reviewer #1)	Results	Why aren't recipients of donor eggs (or sperm) mentioned?	The focus of the KQ6 key question is on the donors rather than the recipients of those donor eggs or sperm.
126.	Jessica Goldstein (Public Reviewer #1)	Results	Not certain how the key points were determined, but the evidence for vitamin E or zinc supplementation does not really make sense from a clinical perspective	This finding was supported by a good-quality systematic review of 4 small RCTs and so although the SOE was rated as low we do retain our discussion of this finding.
127.	Jessica Goldstein (Public Reviewer #1)	Results	The authors characterized publication bias as completed trials registered on ClinicalTrials.Gov not yet published. They infer that even if all of those studies had failed to find a difference between subjects and controls, they would not have swayed the conclusions of their respective sections. The papers cited throughout the document were not all trials that would have been registered on ClinicalTrials.gov. Publication bias can be defined more broadly. There could be many more studies that were performed but found no difference between subjects and controls that never saw the light of day (perhaps beyond abstract submission). This is a limitation of all of the infertility literature and I'd encourage the authors to think a bit more broadly about their definition.	We agree that publication bias can be defined more broadly. We have inserted in to our methods text highlighting the limitations of our approach.

#	Commentator & Affiliation	Section	Comment	Response
128.	Jessica Goldstein (Public Reviewer #1)	Results	The omission of functional hypothalamic amenorrhea is glaring in a document that is 164 pages long.	This was outside of the scope of the key questions prioritized by the stakeholders and therefore was not included within the systematic review.
129.	Jessica Goldstein (Public Reviewer #1)	Results	One area of disagreement with the authors' interpretation of the literature is metformin versus clomiphene citrate in treatment of women with polycystic ovary syndrome (PCOS). The authors state (page vi): "There was low SOE that there is no difference between clomiphene and metformin as primary therapy." The outcome to which this sentence is referring is not clear. If it is referring to the outcome of live birth, then PPCOS1 demonstrates that clomiphene citrate alone is superior to metformin alone. The other RCT included in this analysis (Morin-Papunen, 2012) wasn't really a head-to-head trial of metformin versus clomiphene citrate as monotherapies because both therapies weren't started at the same time.	We agree that PPCOS1 showed superiority of clomiphene to metformin, but the total weight of evidence, including two systematic reviews, lead to the conclusion that the overall difference is not significant.
130.	Anonymous (Public Reviewer #2)	Results	Figure 2, Page 18 - footnote is missing (a)	We have added in the missing footnote.
131.	Anonymous (Public Reviewer #2)	Results	Table 10 is confusing - Results intervention and Results comparator columns are confusing. from where did you get 95%CI for intervention and comparator for study number126 you can add N to the "intervention and comparator" columns	In order to be 508 compliant we are unable to have merged headings in our table – we therefore need to clarify that the specific column is results for the intervention arm and then a second column is results for the comparator arm. We have retained the structure of the table.

#	Commentator & Affiliation	Section	Comment	Response
132.	Louis DePaolo (Public Reviewer #3)	Results	Unfortunately, the cutoff date for the systematic review was September 15, 2017 which precluded inclusion of an important study published by the NICHD's Reproductive Medicine Network in the New England Journal of Medicine approximately one week later on September 24, 2015. The paper entitled, "Letrozole, Gonadotropin or Clomiphene for Unexplained Infertility', contained results of the 900 patient AMIGOS study. It is felt that the study's findings are important enough to include in this review rather than delay until the next review. Among the most important findings were that letrozole and clomiphene resulted in fewer multiple gestations than gonadotropin therapy, and that letrozole, but not clomiphene resulted in fewer live births compared to gonadotropin therapy. The authors concluded that in contrast to women with PCOS where letrozole was found superior to clomiphene for live births, clomiphene may be the drug of choice for ovulation induction and live birth outcomes in women with unexplained infertility.	This article was included in the search update for our final report and is now included in KQ3 findings.
133.	Peer Reviewer #1	Discussion/ Conclusion	Sufficient	Thank you. No response needed.
134.	Peer Reviewer #2	Discussion/ Conclusion	One of the strengths of the document is the fact that the results are compared with NICE and ASRM guidelines as well and there is a good discussion of the differences/issues. It also covers a largely academic discussion of the limitations of the data and how research in this area could be improved. I believe that the challenge with this document is that it attempts to cover the entire specialty and is quite overwhelming to the readership. What is the intended audience? If this is directed towards clinicians, it needs substantial revision to make the clinical recommendations clear. In areas where the evidence is insufficient – then clearly state that there is insufficient evidence to recommend X,Y, Z approach and good clinical practices should be employed.	We understand that the full report is complex and dense and we have attempted to increase the readability of the report in the final version. The incorporation of the evidence from this systematic review however in to clinical guidance is outside of the scope of the EPC program.

#	Commentator & Affiliation	Section	Comment	Response
135.	Peer Reviewer #5	Discussion/ Conclusion	1. Appreciate discussion of issues on page 135 2. Implications are clearly stated 3. Yes,there are specific guidelines that can be used	Thank you. No response needed.
136.	Peer Reviewer #6	Discussion/ Conclusion	The report makes a case for live births as the proper outcome for subfertilty trials even suggesting that various data bases and government agencies change their approach. This is done rather subtlety and I think it needs to be emphasized. It is amazing how many resources are spent on subfertility evaluation and treatment with such a weak evidence base and the rarity of using a meaningful outcome. This needs to be a major message of the ER.	We agree with the importance of this point and we have highlighted the importance of this decision within the implications for policy making section within the discussion.
137.	TEP Reviewer #1	Discussion/ Conclusion	I think the authors nicely summarize the limitations and the research directions seem appropriate.	Thank you. No response needed.
138.	TEP Reviewer #2	Discussion/ Conclusion	Why are the authors equating race/ethnicity with socioeconomic status (ES-17). I am not sure that this statement is uniformly true. If it is, the statement needs to be referenced. Although infertility diagnosis various across countries and within the US, the rationale may not totally be ethnicity but rather the differences in what insurance pays for (like MA, which pays for infertility treatment).	We have revised to clarify that race/ethnicity and socioeconomic status are two different, but often interrelated, concepts. We include data from NASS which suggests that African-American women have different outcomes compared to other racial groups.
139.	TEP Reviewer #2	Discussion/ Conclusion	It might be better to just discuss socioeconomic factors and I think this discussion would also benefit from noting the differences in payment by insurance by state.	Discussion of differences in payment by state is outside of the scope of this report.
140.	TEP Reviewer #3	Discussion/ Conclusion	On page 28, line 17-18 - it's not clear what is meant by live birth per couple? Does that refer to cumulative rate of live birth? If so, would be clearer to state that.	We have clarified that we mean cumulative live birth rate per couple.

#	Commentator & Affiliation	Section	Comment	Response
141.	TEP Reviewer #3	Discussion/ Conclusion	Page 30, line 22. CDC and SART maintain separate databases. Please do not refer to them as a single database. Reporting of ART data to CDC (not SART) is required by law. CDC collects data via NASS (described on page 46). CDC allows SART member clinics to report their data through SART. However, not all clinics are SART members. CDC collects data from SART and non-SART member clinics. Also, SART collects variables that CDC does not. This needs to be corrected throughout the document. There are no studies that use a CDC/SART database. They use one or the other.	We have expanded the description of the national registry of ART procedures to clarify these points, and have revised throughout.
142.	TEP Reviewer #4	Discussion/ Conclusion	The discussion is well written; the comparison of the results of the current analysis with two different sets of guidelines/recommendations from NICE and ASRM is helpful (with references to specific tables)	Thank you. No response needed.
143.	Jessica Goldstein (Public Reviewer #1)	Discussion/ Conclusion	There was a strange sentence in the overall conclusions ("Although the risk of some cancers is increased in women with some infertility diagnoses…"). I didn't find any part of the text that supports this.	We have removed this key message.
144.	Peer Reviewer #5	Clarity/ Usability	Main objectives are clear and conclusions are in alignment with stated goals. I don't believe there are new findings, but a confirmation of beliefs	Thank you. No response needed.
145.	TEP Reviewer #1	Clarity/ Usability	Yes this is well-organized. the conclusions are more helpful in defining areas where we do not have enough knowledge than they are in changing practice in my opinion. I do not think there are many new insights beyond what the specialist in REI would already know. This is due to a lack of studies, and is not the fault of the authors	Thank you. No response needed.

#	Commentator & Affiliation	Section	Comment	Response
146.	TEP Reviewer #2	Clarity/ Usability	The publication is a good first start at looking at a difficult problem in the US, which is how to optimize care for these patients with different diagnosis. The problem is that the paper ignores the fact that many of the ex US publications may or may not reflect practices that are applicable to the US. Outside the US, governmental regulation significantly restricts what can and cannot be done in many countries. Not acknowledging that may lead to conclusions that are not relevant to the US.	We agree that studies based outside of the US may be limited in their applicability to the US setting. Within the applicability section of our discussion we highlight some of the limitations of the evidence base given these potential geographical differences.
147.	TEP Reviewer #3	Clarity/ Usability	Report is well organized and the structure is easy to follow. After reading the evidence summary, I was a bit confused when I started reading the full report. I wasn't aware the first part was just a summary (especially since references were listed). Might be good to make the distinction more apparent.	Although the full report contains both the evidence summary and then the main report, we understand that some readers may only refer to the evidence summary. We therefore have structured this summary to be a stand alone document – but then have the full report to provide additional findings and details.
148.	Peer Reviewer #6	Clarity/ Usability	Please see above. In my opinion they are poised to make stronger arguments about definition of fertility problems (subfertility vs infertility) and proper outcomes (live births) that cd be catalyst to revolutionize investigations in this domain. They seem to downplay the impact this ER cd have and I think they need to 'toot this horn' with more enthusiasm and vigor.	As suggested we have highlighted this recommendation and provided additional citations from other groups.
149.	TEP Reviewer #4	Clarity/ Usability	There are 12 major conclusion from the paper stated in the summary abstract (mentioned in comments below):	Concerns addressed individually below as needed.
150.	TEP Reviewer #4	Clarity/ Usability	1. In PCOS patients, Letrozole compared to clomiphene results in higher live birth rates while reducing multiple births - highly relevant to policy and practice decision making	Thank you. No response needed.

#	Commentator & Affiliation	Section	Comment	Response
151.	TEP Reviewer #4	Clarity/ Usability	2. No Difference between Clomiphene and Metformin as primary therapy (It seems that Table 10 does not support this conclusion, from the large RMN study done by Legro). This in my opinion would be a poor policy or practice decision.	The total weight of the evidence, including other studies and systematic reviews, suggests no significant difference, although there is some concern about reporting bias.
152.	TEP Reviewer #4	Clarity/ Usability	3. Live birth rates are not different between laparoscopic ovarian drilling and oral agents. Moderately relevant for policy and decision making - LOD has fallen out of favor with the increased use of IVF for PCOS patients - but i think it is good to refocus what the literature says about LOD (and there are many PCOS patients that cannot afford IVF)	Thank you. No response needed.
153.	TEP Reviewer #4	Clarity/ Usability	4. For unexplained infertility patients, no difference between letrozole and anastrazole for the outcome of ectopic pregnancy. Low clinical or policy relevance - anastrazole is rarely used for ovulation induction and the outcome of ectopic pregnancy is rare. This is a weird finding that may not be that generalizable	Our findings highlight interventions and outcomes which have been prioritized by our TEP and key informants.
154.	TEP Reviewer #4	Clarity/ Usability	5. Time to pregnancy was shorter with immediate IVF vs COH/IUI. This is highly relevant from both policy and decision making, especially for older women.	Thank you. No response needed.
155.	TEP Reviewer #4	Clarity/ Usability	6. No difference in live birth rate between ICSI and IMSI - not clinically relevant as IMSI is not really used.	Our report summarizes the available evidence of treatments which were determined by the TEP to be of clinical interest.
156.	TEP Reviewer #4	Clarity/ Usability	7. Lower OHSS rate in donors using GnRH trigger - well known, but still clinically relevant from policy and practice decision making	Thank you. No response needed.
157.	TEP Reviewer #4	Clarity/ Usability	8. Lower live birth rates in African Americans - also well known but still highly relevant	Thank you. No response needed.
158.	TEP Reviewer #4	Clarity/ Usability	9. Slightly lower live birth rates with single vs multiple embryo transfer - well known and addressed with SART Guidelines - still very relevant and already implemented into policy and practice decision making	Thank you. No response needed.

#	Commentator & Affiliation	Section	Comment	Response
159.	TEP Reviewer #4	Clarity/ Usability	10. Maternal cancer information is highly relevant from a public perspective also.	Thank you. No response needed.
160.	TEP Reviewer #4	Clarity/ Usability	11. Neurodevelopmental disorders following ICSI - emphasize low SOE - this is known but obviously not understood, and represented an area of more focused research.	Thank you. No response needed.
161.	TEP Reviewer #4	Clarity/ Usability	12. No difference between differing adjunct treatments (should probably define "adjunct treatments" even in the summary abstract) used in combination with oral agents and IUI in live birth rates, miscarriage rates, OHSS in unexplained infertility patients. This statement won't be understood without a definition accompanying "adjunct treatments".	We have added the statement : “Throughout this report, we use the term “adjunct treatments” to refer to interventions performed within a major treatment category (for example, comparison of metformin to placebo as pretreatment in women with PCOS undergoing IVF)” to the Treatment Strategies section of the introduction.