Comparative Effectiveness Review Disposition of Comments Report

Research Review Title: Anxiety in Children

Draft review available for public comment from February 9, 2017 to March 8, 2017.

Research Review Citation: Wang Z, Whiteside S, Sim L, Farah W, Morrow A, Alsawas M, Barrionuevo Moreno P, Tello M, Asi N, Beuschel B, Daraz L, Almasri J, Zaiem F, Gunjal S, Larrea Mantilla L, Ponce Ponte O, LeBlanc A, Prokop LJ, Murad MH. Anxiety in Children. Comparative Effectiveness Review No. 192. (Prepared by the Mayo Clinic Evidence-based Practice Center under Contract No. 290-2015-00013-I.) AHRQ Publication No. 17-EHC023-EF. Rockville, MD: Agency for Healthcare Research and Quality; August 2017. www.effectivehealthcare.ahrq.gov/reports/final.cfm.
DOI: https://doi.org/10.23970/AHRQEPCCER192.

Comments to 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 in draft form for public comment for a 4-week period. Comments can be submitted via the EHC Program Web site, mail or E-mail. At the conclusion of the public comment period, authors use the commentators' submissions and comments to revise the draft comparative effectiveness research review.

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

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

.

Source: https://effectivehealthcare.ahrq.gov/topics/anxiety-children/research-2017/
Published Online: August 31, 2017

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	Abstract	Results: I would be inclined to put the sentence "No individual medication had evidence…(lines 35-37)" after the sentences about effectiveness rather than before. Results: It is a bit confusing to call out CBT as the modality that leads to anxiety reduction (line 47) and then in the next sentence say it is no different than other modalities. Results (pg 7, line 5) CBT was more effective than what?	We moved "No individual medication .." to the sentences behind the effectiveness. For the other comments, we think they are clear in the revision now.
Public Reviewer # 1 **[Raquel Halfond, American Psychological Association]**	Abstract	Abstract: Results • 'compared to placebo' repeated twice in third sentence. • For the sentence: Subgroup analyses suggested that CBT was more effective in reducing primary anxiety symptoms in patients with ADHD (compared to autism), when it was provided in school setting (compared to outpatient or mental health settings) and when it was associated with exposure therapy.	We corrected these.
Peer Reviewer #1	Key Messages	pg 8, line 16) (Would say that medications "may" result in adverse effects	added
Peer Reviewer #1	Introduction	Background (pg 11, line 12) To call anxiety disorders "unremitting" and "unrelenting" seems a bit overbearing and not supported by the evidence. For example, a large number of kids tend to outgrow separation anxiety disorder, although many don't (Silove et al., AJP 2015).	Thank you for the comment. We change it to "impairing and often continue into adulthood."
Peer Reviewer #1	Introduction	Background (pg 11, line 20) It is a little surprising not to see the American Academy of Child & Adolescent Psychiatry not mentioned here among the others. They publish specific guidelines on the treatment of child anxiety disorders. Their Practice Parameters on anxiety (www.aacap.org) would seem very relevant resource here. Indeed the "treatment guidelines" that the authors cite (line 37) are not actually real guidelines but rather recommendations based on a single study.	AACAP is one of the partners for this study. AACAP guideline is cited as Ref 19. The treatment guidelines are referenced with 3 guidelines (AACAP, NICE, and British Columbia Medical Services Commission).

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	Introduction	Background (pg 12, line 22) Another reason PTSD and OCD might not be considered here is that, at least according to the DSM-5, they are not anxiety disorders.	We listed the reason in the introduction.
Peer Reviewer #1	Introduction	Scope and Key Questions Table 1. Recognizing that it is difficult to categorize, many people will likely take issue with what is deemed CBT and non CBT. EMDR and mindfulness have quite a bit of CBT elements, for example.	The treatment categories in the initial draft were preliminary. We have revised the categories and no longer have a non-CBT category.
TEP 1	Introduction	p. 6 lines 37-40 “compared to placebo” is repeated at the beginning and end of sentence	Thank you. We corrected this.
TEP 1	Introduction	p. 6 lines 37-40 p.6 results: I wonder if this section is too unclear because of the missing context. For example, comparisons specifying medication or CBT vary by disorder but that is lost in the summarized results. Likewise, within the tables, listing diagnosis(es) may help with clarity.	The vast majority of studies combine diagnoses. As such our primary analyses used studies that combine diagnoses. When there is a sufficient sample of studies we examine disorders separately.
Peer Reviewer #2	Introduction	The topic is introduced clearly. I recommend revising the statement of prevalence rates for anxiety disorders in youth. Citing a specific rate (8%) is not well supported, a fact that is made clear in the subsequent sentences.	Here we used NIMH estimates.
Peer Reviewer #2	Introduction	Growing evidence for family factors impacting outcomes for child anxiety treatments such as family conflict and family accommodation suggest they would have been useful to include in studies or in presentation of background and future research.	We added family dysfunction/stressor in the factors that may affect treatments. We had this as a subgroup analysis. But we were unable to evaluate it as few studies reported this. We reported it as a limitation and future research needs (page 46).
Peer Reviewer #2	Introduction	Using the term 'withdrawal' in the context of treatment (in particular medication) studies may cause confusion. Readers may interpret this to mean 'withdrawal syndrome'. I suggest considering an alternative term for cessation of treatment (such as cessation of treatment).	Excellent observation. We used drop-outs in the revision.
Peer Reviewer #2	Introduction	I would recommend adding parent based therapies to the list of available treatments including SPACE a treayfocused on reducing family accommodation.	We have added a category for parent based treatments. Relevant studies can be found in Appendix Table E8-9.

Commentator & Affiliation	Section	Comment	Response
TEP 3	Introduction	The introduction is well done. A minor point is the introduction did not mention the black box warning for antidepressants which is a concern for parents and other stakeholders in terms of adverse events. Also, several of the anxiety medication studies included in this analysis were done during the time and I believe one prematurely stopped due to recruitment challenges.	We added a discussion of the black box warning on page 46 of the discussion.
Peer Reviewer # 3	Introduction	The introduction to the report appropriately frames the problem of interest, but in this section the authors promise a much more ambitious report than they deliver. In particular, there is strong emphasis on the need for subgroup analyses within this literature and the suggestion that the report authors will be in contact with trialists to obtain data for these analyses, when published reports do not provide necessary information. As discussed in later sections of our review, the subgroup analyses were quite weak and did not appear to make use of available data in the field. There is a similar mis-match on other issues, when the authors promise a report that will exceed the scope and detail provided in other products in the field and then fail to deliver.	We thank the reviewers' comments. We understand the importance of subgroup analyses. However, we couldn't evaluate these due to limited the number of studies with available information/combinable data. In this revision, we added some subgroup analyses. We also emphasized the inability to conduct subgroup analyses in the discussion.
Peer Reviewer # 4	Introduction	The writing is very good and clear.	Thank you.
Peer Reviewer # 1	Methods	(pg 16, line 53) While industry funding was not considered an automatic indicator of high bias, was it considered to be an indicator of any bias?	Yes. Funding source is a part of quality appraisal (listed under "other sources of bias for RCTs and possible conflicts of interest for observational studies). It is not an automatic reason to suspect bias. Reviewers can make a judgment on a case by case basis as to whether it represent a major concern.
TEP 1	Methods	The inclusion/exclusion criteria and process might be elaborated on for greater clarity.	This inclusion/exclusion criteria were listed in the study protocol and used in the study process. We felt it is clear and may be inappropriate to make changes now.

Commentator & Affiliation	Section	Comment	Response
TEP 1	Methods	With regard to the SOE ratings that were made, it would be helpful for the authors to include greater detail on the rating process (how many raters were there, what happened when there were disagreements, were any reliability calculations conducted etc…). Further, given the complexity of the SOE procedures and their heavy use in examining the data/studies, it would be helpful to include a table that presents the starting point for each and then lists the deductions that were made to result in the final rating.	We added more details to the methods section. Particularly, to emphasize that rating was done via consensus of team members with expertise in guideline methodology (and not by independent pairs). The starting point in studies of intervention is high for RCTs and low for observational studies. The requested table is added to the appendix.
TEP 1	Methods	The interpretation of the SMDs and RRs that occurs on page 56 should be moved forward to the beginning of the paper where this is introduced (page 17) so that readers can be examining and interpreting the findings as they read.	This guide for conversion logically fits in the applicability section. However, as reviewer recommended, we added to the methods section further explanation (earlier in the report).
Peer Reviewer # 2	Methods	Inclusion/exclusion criteria are reasonable and search strategy is appropriate. Measures are appropriate. Methods are current and appropriate.	Thank you.
TEP 2	Methods	Inclusion and exclusion criteria were justifiable. Search strategies were logical and explicitly stated. Definitions and diagnostic criteria for outcomes measures were appropriate. Please see attached for one question about statistical methods in terms of subgroup analyses.	Thank you.

Commentator & Affiliation	Section	Comment	Response
TEP 2	Methods	p. 15 CBT vs Drugs: Although it is common to combine social anxiety disorder, GAD, and separation anxiety disorder in treatment studies, I do not think it is an equivalent test of CBT or a medication to compare results from a sample that only includes patients with one of these presenting problems against a sample that includes a mixture of these as primary problems. Specifically in this section, among these studies, I think sertraline had a "harder test" than fluoxetine because I think CBT for social anxiety disorder is a very potent intervention but CBT for a sample with a mixture of primary anxiety disorders is a tougher test for CBT and therefore more likely to not separate from the medication arm. Therefore, I think a caveat is needed when interpreting results across studies that had different primary presenting problems.	Excellent point. We addressed this issue by confining the primary analyses to studies that included mixed samples. The subgroup analyses examined disorder specific studies.
TEP 3	Methods	I thought all very clear and justifiable. Although I did not understand why observational studies and case series for adverse events were included since most of the studies included were RCTs.	RCTs are not particularly designed for detecting adverse events (short followup, small number of patients). This is particularly a problem for rare adverse events. By including observational studies, we capture a better assessment of adverse effects.
TEP 3	Methods	I thought the sections explaining the definitions of SOE was well done and helpful as one went on to the results.	Thank you.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer # 3	Methods	Selection of studies. We have serious concerns about the methods used for this investigation on two fronts. First, insufficient detail was provided on key methodological decisions (e.g., how SOE of individual studies was aggregated at a category level). Second, there appear to be several errors in either the methods themselves or the implementation of the methods. A list of our most substantial concerns follows: The authors provide extensive and adequate information on there search strategy for studies. However, little information or justification is provided for study inclusion / exclusion. For example, some foreign-language studies seem to be included, but not others. Authors were entirely unclear on how they handled split publication and papers stemming from the same dataset. Authors also did not sufficiently address the methodological differences between studies that included only pharmacological agents, only psychosocial interventions, or included both, other than to note that psychosocial-only studies tended to have information from multiple informants. These studies vary dramatically in the literature in terms of their outcomes assessed (e.g., little information on AEs in psychological only reports) and the clarity of the definition of the independent variables (agent X at dose X versus "a collection of CBT techniques" and at wide range of sessions). W also had difficulty following the evaluation process that determined if the methods of studies were strong enough to remain in the evidence pool (e.g., sample sizes per cell [e.g., Mancini, 1999], inclusion/exclusion criteria, operationalization and control of experimental conditions, language of origin; see chapter 15 of AHRQ Method Guidelines, 2014). The authors also appear to have adopted an open door policy on study N -- investigations with a comparison group N = 3 are included in the report. In combination with our next point, this was a difficult decision to interpret.	To clarify some of the specific issues brought up by reviewer: 1) there is no such thing as “how is SOE of individual studies is aggregated”. SOE is never made for individual studies, 2) we are very clear in that we did not include foreign language studies in analysis but we did evaluate them and accounted for them, 3) we did not have restrictions on study size (very clear in our methods), 4) studies with multiple publications are treated as a single study in analysis to prevent participant overlap (basic statistical concept), 5) we did not treat studies differently in terms of appraisal based on the intervention. We appreciate the desire for more details on the methods and in the revised manuscript we tried to add some additional information. However, we followed the reporting guidelines (PRISMA). And other reviewers were quite satisfied with the details of the methods. In addition, there are additional details in the study protocol. Making the report longer will reduce its uptake and helpfulness.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer # 3	Methods	Methods for assessing study level and aggregating category level strength of evidence (SOE). The EPC methods guide on assessing the strength of evidence was cited in the report method section; however, after reviewing the contents of that extensive document how the guidelines were applied to this particular exploration remained unclear. We would recommend providing additional details within the methods section explaining (a) how studies were evaluated at the individual level, including clear guidelines for exclusion, (b) how studies were grouped, including clear rationale for using outcome measures as the grouping mechanism , and (c) how results may have been impacted by these decisions (e.g., is it possible that the heterogeneity within group methodologies lead to statistical imprecision?). The Grading the Strength of Evidence (SOE) section on page 8/18 indicates that RCTs start out as having “high” SOE, while observational studies start out “low”; additional criteria then follow as rationales for further alterations to SOE ratings, resulting in categorizationsof “insufficient”, “low”, “moderate”, and “high” SOE. This is helpful for improving the reader’s understanding of how studies moved between categories; however, much of the explanations within categories is quite vague (e.g., “arbitrary” cutoffs, “moved down one or two levels” based on attribute severity). It is also unclear which methodological or statistical features would illustrate an exemplary member of each SOE category, so the reader does not have a reference point for the target quality of each category. Furthermore, it is unclear how the SOE ratings were summated within study groupings. Without additional operationalizations of how SOE was assigned, it is difficult to follow how conclusions were drawn from the literature. Please incorporate responses to the following in order to clarify this section:	See above answer. SOE is never determined for individual studies. We understand the reviewer discomfort about not having an algorithmic way (or quantitative) to determine SOE. SOE after all, is a judgment. In the revised manuscript, we added some more details to the methods section; nevertheless, SOE remains a judgment. We added supplemental tables that have more explanation.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer # 3	Methods	At what level is the SOE is assigned , individual study or group level? If the former, please clarify the explanation of how the SOE ratings were summarized within study groups. In particular, please address how you adjusted the evidence due to notably different methodological rigor of individual investigations that were grouped together. Differences in methodology introduces heterogeneity within the larger groups of studies. It is possible that this variability could be driving results, given the apparent "open door" policy toward trial inclusion. It would be helpful to assign SOE categories at the individual level and document those rulings within the tables to clarify how study quality at the individual level translated to SOE rulings at the group level.	See above. SOE can only be assigned at a group of studies (body of evidence). We added a sentence to the methods section to emphasize this issue. We disagree with reviewer that we have an “open door” for study inclusion based on size. This is a common approach of evidence synthesis and restricting studies based on size is a mistake (done only sometimes for feasibility).
Peer Reviewer # 3	Methods	Please provide clearer operationalization of the rules used to alter SOE categorizations. Particularly, it would be helpful to see more specific guidelines for when studies were penalized one versus two levels within each of the SOE domains.	Same answer as above. The section of SOE has sufficient details.
Peer Reviewer # 3	Methods	Please remark on which methodological limitations were most weighed most heavily in drawing SOE conclusions (e.g., it seems like concordance among informants was weighed heavily; however, we cannot identify throughout the manuscript mention of this characteristic which makes it difficult to draw the same conclusion).	We believe reviewers are mixing the two constructs (risk of bias and SOE). In terms of ROB, we explicitly stated which elements were emphasized. For example, for RCTs, we said “A judgment of overall risk of bias across the various domains was made focusing on random allocation, allocation concealment and blinding (high risk of bias in any of these domains led to a high overall rating). We did not consider industry funding as an automatic indicator of high risk of bias. ”. Secondly, concordance among individuals was NOT considered in the SOE because we considered results based on each informant to be a separate independent outcome. This is clarified in the methods section.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer # 3	Methods	Definitions of outcomes. As noted above, the authors seem to group studies and weigh SOE in part on whether results were shown across informant domains. We concur that interventions that show evidence of efficacy across multiple methods are especially impressive. However, it is not clear how this assessment was made (see comments on aggregating SOE above).	Concordance among individuals was not considered in the SOE because we considered results based on each informant to be a separate independent outcome. This is clarified in the methods section.
Peer Reviewer # 3	Methods	Adverse event reporting. This section addresses KQ2 that focuses on harms and safety concerns related to treatments. Studies are required to report any adverse events that occur during study participation; however, not all AEs are protocol related nor are they equally severe in nature. Clarifying the nature of the AEs reported would assist the consumer. In the spirit of making the final report "patient-centered" as dictated in the guidelines, it may be useful to include greater detail in this section to assist with defining AE categories and whether or not they were protocol related (e.g., cold/infection/allergies). Additionally, when withdrawals are listed, does that refer to dropping out of the treatment or experiencing withdrawal symptoms? Please clarify, particularly if you are using both definitions at various times. Furthermore, this section lists harms associated with CBT but no other psychotherapeutic categories. Please expand on this in text.	We agree with the reviewers on this. However, studies seldom reported AE severity or protocol related. In addition, most of the time, it is unclear how they rated severity. We believe it's best to handle this without addressing severity as there are so many "unknowns". For withdrawals, we reported overall withdrawals and withdrawals due to adverse events. The term, "withdrawals", was also changed to "dropouts". Per Key Question 2, harms were assessed for all treatments.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer # 3	Methods	Definition and classification of interventions. Our final major methodological concern deals with classification of interventions, particularly inconsistencies and errors in the classification of psychosocial interventions. The authors do not provide a clear definition of "CBT" or justification for the aggregation of other wildly different intervention models (attention modification vs. psychoanalysis!) in the umbrella "other psychosocial" category. This methodological flaw is best illustrated by a description of the specific studies involved, and the remainder of our comments on this topic are included in our critique of the results section below.	We have revised the categories and no longer have a non-CBT category. We no longer have a categorical CBT vs. “other” comparison as there is not a sufficient sample of homogenous “other interventions” to examine.
Peer Reviewer # 4	Methods	So sophisticated. It didn't appear that any trials were excluded, especially very small studies.	Thank you.
Peer Reviewer # 1	Results	Table 6. Some studies appear to be missing in the analyses. For example, the CAMS study (ref 47) that compared placebo, sertraline, CBT, and combination is not referenced in the sertraline vs pill placebo areas despite the fact that it did compare sertraline vs placebo.	The CAMS was included in the original draft. In this revision, we cited all relevant references whenever the CAMS was listed.
Peer Reviewer # 1	Results	Later (pg 36, line 42) the reference is only for the follow-up phase but not the initial trial. All this just raises a little doubt about whether the right references are used in the right places.	We corrected this.
Peer Reviewer # 1	Results	(Pg 27, line 7) Unless I missed it, explaining the difference between placebo and pill placebo might be useful.	We re-categorized the control group to 1) pill placebo, 2) waitlisting or no treatment, 3) attention control or treatment as usual
Peer Reviewer # 1	Results	For many of the tables, is it truly necessary to have in the “outcome” column all the names of each individual rating scale. It makes the tables quite unwieldy.	We removed the scale names.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer # 1	Results	It might make sense to have a table summarizing the evidence for each individual types of therapy (psychoanalysis, etc) similar to what was done for individual medications. This would alleviate some of the problem with lumping together some many varied types of approaches. (pg 46, line 38) There continues to be enormous controversy over the possible induction of suicidal behaviors from SSRI, yet this review seems to ignore this important issue completely. Indeed, the section related to SSRIs doesn't even have a column for it.	CBT is defined as attempts to change cognition and behavior generally consisting of some combination of cognitive restructuring, relaxation training, and exposure therapy. CBT involves working with child directly and some degree of parent involvement. Studies that did not met this definition were included in the appendix table and not in the meta-analyses. For suicidal behaviors, we deemed it as adverse events and summarized in results related to Key Question 2. However, only 3 studies reported suicidal behaviors (March 2007, Renaud 1999, and Walkup 2008). We added this in the results.
TEP 1	Results	A number of findings presented need to be more clearly articulated or corrected as they do not match the findings depicted in the corresponding tables or are unclearly explained. These instances are detailed below: Page 19, line 21: "3 studies reported AEs from psychotherapy" – it is unclear which 3 studies are being referenced here as this statement does not correspond with the findings in the tables. Rather, the only psychotherapy findings that appear to show AEs are those where psychotherapy was combined with a drug (CBT+ sertraline), and I do not think the studies disentangled which aspect led to the AE which should also be noted.	We extensively revised the manuscript and added the reference to those numbers. The studies compared CBT to CBT and medications. If the studies reported any AEs in CBT, we counted it. We also added citations.
TEP 1	Results	Page 20, line 11: should specify that it is clinician-reported improved function where support was found. These specifications (who the reporter was) should be made throughout this section (presently, it is in some places specified but not in others). The authors should also clearly state that across the findings in Table 5, there is some support from clinician-rated measures and indices, however parent and child reported measures did not consistently support the use of any of the drugs examined. This is a critical distinction.	We were unable to separate it as most of the studies were not clear and it was not in our study protocol.

Commentator & Affiliation	Section	Comment	Response
TEP 1	Results	Page 25, key points: the authors might clarify that both of the RCTs comparing CBT to drugs used SSRIs (not all readers will have in depth knowledge of different drug classes).	Changed.
TEP 1	Results	Page 27, line 33: please define “sham Psychotherapy”	We deleted this word.
TEP 1	Results	Page 36, key points: clarification should be made that these points are being drawn from 1 study in each case.	We added this point.
TEP 1	Results	Page 38 (and in analyses): please clarify what differentiates a mental health clinic from an outpatient clinic.	We clarified now that outpatient clinic as outpatient primary care.
TEP 1	Results	Page 38, Comorbidity: the findings reported in the table (parent and clinician reports are significant) do not match what is reported here (child and parent).	Changed.
TEP 1	Results	Page 38, Treatment Settings: the findings reported in the table (child and clinician versions as significant) do not match what is reported here (child only).	Changed.
TEP 1	Results	Page 38, Diagnosis: the findings reported in the table (parent and clinician reports are significant) do not match what is reported here (clinician only). The subgroup analysis section (page 38) requires additional details and should be explicit about what was examined and what was found to be significant, as well as what was not significant.	Changed.
TEP 1	Results	Page 39, Key points: the authors should define severity as referenced (the AEs are noted to be “overall not serious”).	We are quite clear about this issue and we have evaluated side effects as “any side effect”, “side effect that leads to withdrawal” and then each specific adverse effect (eg, GI, headache, etc). There were only minor adverse effects reported, we did not identify serious ones (which is consistent with practice).

Commentator & Affiliation	Section	Comment	Response
TEP 1		Page 39, lines 30-35: Several of the findings reported in the tables are not included in this summary and need to be added, including: Venlafaxine was also associated with withdrawals due to AEs, which should be added to the existing list in this section. There are several additional instances where such omissions occur in the text summary on page 39 but are apparent in the tables. These include that: Paroxetine vs. Placebo was associated with Any AEs; SNRIs vs Pill Placebo was associated with AEs related to fatigue-somnolence; and SSRIs vs. Pill Placebo was associated with AEs related to behavior change and AEs related to Cold/infection/allergies.	We try to report the main findings in the summary sections. Details can be found in the texts and tables.
TEP 1	Results	Page 39, lines 35-37: it is described that "CBT was associated with more withdrawals than placebo" but this is not consistent with what is presented in the table (which shows that this occurred for CBT versus Waitlist not placebo).	We changed this.
TEP 1	Results	Table 6 is challenging to follow. At a minimum, it would be helpful to indicate what class of drugs each comparison comes from (e.g., SNRI, SSRI, TCA). Adding a separate table that summarizes the findings broadly might help (e.g., Outcome areas across the top; comparison along the left side; level of support marked in the corresponding boxes).	We added the drug class in the table along with the specific drug name. For the overall conclusions, we summarized as the key points at the beginning of each comparison.
Peer Reviewer # 2	Results	Results are clear and cogent. Key messages are clearly conveyed. Figures and tables are adequate. Results do not address al the aspects of the key questions: Age differences, family burden and contextual differences, etc are stated as clauses of key questions but not addressed in any depth.	Due to the lack of data, we were unable to conduct most of the planned subgroup analyses. We addressed this in the limitations and future research needs.
TEP 2	Results	Results were appropriately detailed. Studies were clearly described. Key message were explicit (and with the one caveat in the attached were applicable. Tables, figure and appendices were adequate (extensive :). I am not aware of any study that was overlooked nor am I concerned that a study was included that should not have been included.	Thank you.

Commentator & Affiliation	Section	Comment	Response
TEP 3	Results	I do not believe any studies were missed. The results section is detailed along with the figures and tables. It is clearly written. It is a great deal of information and takes time to work through it. I am not sure if there is a way to streamline it.	Thank you.
TEP 3	Results	I believe there was a line that stated there were no serious adverse events in the studies that involved medications. I am not sure how that was defined and it might useful for the readers to have more information about that. Also, when talking about discontinue it was not always clear if it was referring to discontinuation for any reason or due to an AE.	The adverse effects were minor for the most part. Nevertheless, we present each specific adverse events so that readers can make these judgments on their own. In the averse effects table, discontinuation due to an AE is explicitly described as such.
Public Reviewer # 1 [Raquel Halfond, American Psychological Association]	Results	For section on other types of psychotherapy vs. placebo, waitlist, treatment as usual: It would be helpful for the reader to include a little more information in the main body of text about why modification of CBT treatment was not classified together with CBT treatments and also more explanation of what single session therapy consisted of.	In the revised manuscript, we made some changes in treatment categories. CBT is now more clearly classified.
Peer Reviewer # 3	Results	In short, no. As discussed previously, there is a serious mismatch between the ambitious aims of the report in the introduction, the impoverished information provided in the methods, and the errors and inconsistencies in the results and tables. Our main concerns in the results section are: (1 and 2) a variety of errors, omissions, and unusual decisions in the categorization of studies in the psychotherapy sections, (3) poor attention to split publication, and (4) inadequate subgroup analyses.	In this revision, we extensively edited the contents. All of the outcomes and categorizations of outcomes/interventions have been double checked and updated. Unfortunately, the reviewers comments are overarching and not specific enough to allow further answers.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer # 3	Results	Critique of “CBT versus other Psychotherapy” results comparison (pp. 34-36). While we agree that the comparative efficacy of different forms of psychotherapy represents an important area for examination (KQ1.c), we feel that the analyses as presented in this section potentially misrepresent the state of the literature base. As currently framed, the authors present a comparison of “CBT versus other psychotherapy.” Our concerns regarding these analyses are twofold. First, we do not feel that it is theoretically meaningful to group all non-CBT psychotherapies into the same category. This is akin to early meta-analyses of psychotherapy that compared categories such as “verbal” and “behavioral” therapies, an approach which has been highly criticized (for a review, see Hunsley & Di Giulio, 2002). Of relevance to the current review, it does not seem appropriate to combine therapies such as attention bias modification (Waters et al., 2014 [147]) with emotional disclosure (Muris et al.,2002 [149]), as these therapies have very different theoretical foundations and empirical bodies of support. Secondly, we feel that some of the specific categorizations of “CBT” and “other psychotherapy” in the current review are not valid. Of the seven studies cited with comparisons between CBT and other psychotherapy, two explicitly compare different forms of CBT (group CBT versus individual CBT, Mannasis et al., 2002 [148]; child CBT versus family CBT, Bodden et al., 2008 [150]), one compares exposure treatment to a “psychotherapy” condition that includes other components of CBT such as an emphasis on behavioral strategies and removal of secondary gains (Millek et al., 1972 [116]), and one included two exposure groups with an adjunctive attention bias modification condition in one group (Waters et al., 2014 [147]). Thus, only three of the seven studies clearly compare variants of CBT to other forms of psychotherapy (Abbasi et al., 2016 [138]; Ollendick et al., 2009 [85]; Muris et al., 2002 [149]). We suggest that this section warrants major modification and that it may be inappropriate to compare CBT (which is itself already a broad category worthy of parsing into various components) to the broad category of “other psychotherapy.” Instead, the authors may narratively comment on the	In the revised manuscript, we made some changes in treatment categories and no longer have a non-CBT category. We no longer have a categorical CBT vs. “other” comparison as there is not a sufficient sample of homogenous “other interventions” to examine.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer # 3	Results	Critique of "Other types of Psychotherapy versus placebo, waitlisting, treatment as usual" comparison (p. 32-34). Per the authors, "Other types of psychotherapies included non CBT treatments (i.e., attention bias modification, modifications of CBT for patients with autism spectrum disorder, single session therapy, and psychoanalysis)" (p. 32). It is unclear from this definition as well as the tables provided by the authors (Table E.4 and Table 10) which conditions were grouped into which categories. Of the 8 studies cited by the authors in "Table 10. Strength of Evidence for other types of psychotherapy versus placebo, waitlisting, treatment as usual," 5 appear to include conditions that are various forms of CBT (Goldbeck & Ellerkamp, 2012 [142]; Hampe et al., 1973 [145]; Miller et al., 1972 [116]; Ollendick et al., 2009 [85]; White et al., 2013 [144]). Given the heterogeneity of techniques and length of sessions used in the studies actually included in the CBT condition in different sections of this report (e.g. Ollendick et al 2009 [85] one-session exposure treatment is included as a CBT condition in the section "CBT versus other psychotherapy") why would these studies specifically be separated and labeled as "other psychotherapy" instead of CBT? Moreover, for studies with more than 2 conditions (Hampe et al., 1973 [145]; Miller et al., 1973 [116]; Ollendick et al., 2009 [85]) it is unclear which conditions were used for the comparison of interest. Relatedly, this section also requires clearer definition of how the authors defined "other psychotherapy" versus "placebo." Furthermore, and these considerations aside, we believe it is not justifiable to combine very different treatments such as attention bias modification (Waters et al., 2013 [141]; Waters et al., 2015 [143]), education support group (Ollendick et al., 2009 [85]), and psychodynamic therapy (Gottken et al., 2014 [146]) in the same "other psychotherapy" category due to the significant theoretical and methodological differences between these therapeutic approaches. Finally, we note that the studies included in "Table 10 Strength of Evidence for other types of psychotherapy versus placebo, waitlisting, treatment as usual" and "Table E.4 "Characteristics of studies comparing other types of psychotherapy versus placebo, waitlisting, treatment as	In the revised manuscript, we made some changes in treatment categories and no longer have a non-CBT category. We no longer have a categorical CBT vs. "other" comparison as there is not a sufficient sample of homogenous "other interventions" to examine.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer # 3	Results	3. Critique on handling of duplicate/split publication We wish to raise some questions regarding how split or replicate publication was handled by the authors. Was this considered? For example, in the section “CBT combined with drugs” and subsection “Sertraline vs. Pill Placebo” (pp. 36-38) the three studies listed in this category (Gonzalez et al., 2015 [46]; Keeton et al., 2013 [40]; Piacentini et al., 2014 [47]) all use the same data from the Child/Adolescent Anxiety Multimodal Study (CAMS), although the original publication which has data on the comparison of interest (CBT+sertraline vs. sertraline) is not cited here (Walkup et al., 2008). Of note, the original Walkup et al. 2008 study is cited in other comparisons of interest (drugs versus placebo, sertraline versus pill placebo, CBT versus sertraline, CBT versus pill placebo, and CBT+Sertraline vs CBT) – why not in the CBT+sertraline vs. sertraline comparison? Walkup, J. T., Albano, A. M., Piacentini, J., Birmaher, B., Compton, S. N., Sherrill, J. T., ... & Iyengar, S. (2008). Cognitive behavioral therapy, sertraline, or a combination in childhood anxiety. New England Journal of Medicine, 359(26), 2753-2766.	We identified 7 publications based on the same CAMS trial. However, not all of the studies reported the same outcomes. The reporting and citation in tables are based on how outcomes were reported.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer # 3	Results	4. Critique on subgroup analysis (pp. 17, 38, 56-57) The authors propose an ambitious set of subgroup analyses in the method section (p. 17); however, it is unclear of those proposed which were actually analyzed, for few of them were actually reported in the results section and/or the summaries seemed incomplete (e.g., comorbid results excluded mention of depression, diagnosis results excluded mention of GAD, Panic, Social Phobia; p. 28/38). The authors state that this "was due to studies not providing sufficient stratified data per subgroup variable" (p. 38). We found this explanation unsatisfactory, as the authors state in the methods section, "When there was missing information, we contacted the authors" (p. 16). Were authors contacted for information that would enable subgroup analyses? If so, did they respond? Additionally, were secondary outcome papers that may have contained additional information needed for subgroup analyses identified and consistently included in the review? Many of the larger clinical trials cited (CAMS, the Coping Cat investigations) have many, many secondary analyses published from these data sets, including papers focusing on predictors and moderators. Please ensure that there is a report of all completed analyses in the results section and clear explanations for why some were incomplete. Implications of the limitations that barred subgroup analyses should be addressed in the discussion section.	We thank the reviewers' comments. We contacted authors for missing data. But few responded. To identify these subgroups, we actually used all relevant information. For example, for the CAMS trial, we found 7 studies. All of the information were extracted and analyzed whenever they are combinable. Subgroup analyses are always ambitious in the protocol, but only those with sufficient data are feasible.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer # 4	Results	The paper is likely to be statistically sophisticates but the results themselves are extremely difficult to review. The overview and summary are the sections most readable. Simple ways of describing outcomes such as effect sizes or NNT would be much better. The only results that are easy to understand are the limitation on the SOE and the SOE, which largely dismiss most all studies. In that sense very small studies and high quality (NIMH funded) are all treated the same. Probably not a helpful approach to the evidence base.	Regarding the quality (NIMH funded or small studies), we believe that the issue is not about who funded the study or the sample size of the study. We used standardized quality appraisal tools (Cochrane Risk Bias tool and Newcastle Ottawa scale) to comprehensive judge the quality and summarized the report.
Peer Reviewer # 1	Discussion	Again, I think there is a problem with how studies are weighted.	Modern meta-analysis does not use factors (other than statistical factors) to weight studies.
TEP 1	Discussion	I believe both the implications and limitations sections would benefit from elaboration. The future research section provides some clear recommendations for moving forward.	Both the implications and the limitations section have been expanded. We added some recommendations to the future research section
Peer Reviewer # 2	Discussion	The discussion is clear and useful. The omission of key topics listed in aims and not addressed (mentioned above) remains a weakness in discussion as well.	We addressed this in the current revision.
TEP 2	Discussion	The implications of the findings are clearly stated and the limitations are appropriately reviewed. The discussion included the important literature. The future research section was clear and could be easily translated into new research. (Now if only funding agencies would agree to support these important studies.)	Thank you.
TEP 3	Discussion	Yes, the implications of the major findings are clearly stated and the limitations are described. Based on the SOE criteria many of the trials seem to be inadequate even though from the perspective of a clinician to have any pediatric trial with a sample greater than 100 seems like a success. It is a sobering reality. If possible it would be helpful if the authors could speak to how they see the challenges to and solutions for to conducting pediatric clinical studies that could move the field forward.	We expanded future research needs.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer # 1 [Raquel Halfond, American Psychological Association]	Discussion	• For the sentence: Exploratory subgroup analyses suggested that CBT was more effective in reducing primary anxiety symptoms in patients with ADHD (compared to autism), when it was provided in school setting (compared to outpatient or mental health settings) and when it was associated with exposure therapy.	We revised this sentence.
Public Reviewer # 1 [Raquel Halfond, American Psychological Association]	Discussion	• Confused by this sentence: A meta-analysis that focused on parent involvement in CBT did not significantly change treatment outcomes.166 Does this intend to mean that the parent involvement did not change treatment outcomes? If so would modify sentence like this: A metaanalysis that focused on parent involvement in CBT found that parental involvement did not significantly change treatment outcomes.166	We revised this sentence.
Peer Reviewer # 3	Discussion	The discussion section has some useful ideas but is thin. In particular, the authors could comment more on the need for consistent reporting of subgroup effects (and suggest the most theoretically important subgroups to include), discussion of development (studies span a VERY wide age range), consideration of functional improvement in outcome assessment, and long-term comparative efficacy of interventions.	Thank you for these suggestions, we have elaborated on these points in the discussion
Peer Reviewer # 4	Discussion	I think the evaluation of CBT is over stated. The value of medication is not nuanced. Larger definitive studies with multiple publications that examine a myriad of findings are not highlighted. All the recommendation appear on the surface appear to be reasonable except they likely are not feasible.	We attempted to balance between CBT and medications as much as possible and remained consistent with the data. Nuanced recommendations will be generated by guideline panelists and not by the systematic review. We conducted multiple subgroup analyses as much as data were available to help in providing more granular inferences.

Commentator & Affiliation	Section	Comment	Response
TEP 1	Clarity and Usability	I do not believe the main points necessarily add new information or understanding, as noted by the authors, the findings are quite consistent with previous reviews in this domain. Given this, it may be helpful for the authors to revisit some of the differences observed in this large scale review. For example, it is remarkable that with drug studies, clinicians tended to report improvements while parents and children did not consistently do so, whereas with CBT studies, improvements seemed to be more commonly reported across all reporters.	We added this point to the discussion. We thank reviewer.
Peer Reviewer # 2	Clarity and Usability	Report is well structured and organized. Points are made clearly. Conclusions are relevant to practice and policy. Not a lot of new information but good review of current data.	Thank you.
TEP 2	Clarity and Usability	The report is well organized and structured. The main points are clearly presented. The conclusions have some policy implications (i.e., CBT and/or SRIs should be the treatment of choice of child anxiety disorders). The report has less to offer in terms of guiding practice decisions. The limits on the subgroup analyses and the limits on what we know (don't know) about specific meds (SSRI) and specific components of CBT severely limit the utility of this report for guiding treatment decisions for individuals.	Thank you for the comments. We addressed this (unable to conduct subgroup analyses) as a main limitation and identified as future research needs.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer # 3	Clarity and Usability	The draft has significant conceptual errors as noted previously, and it is difficult to find key methodological details. The writing also needs to be cleaned up considerably (e.g., missing words, use of SRI instead of SSRI, incorrect punctuation, order of text not matching the table order, inconsistent abbreviations). The impact of this report would be significantly strengthened by subgroup analyses, as proposed in the report's own introduction. However, these analyses were not conducted, despite some missed opportunities to gather this information in the literature (e.g., by coding characteristics of included reports [studies by age group, studies by ethnic representation], by considering secondary papers from trial datasets examining predictors and moderators). Conclusions of the report are broadly in line with existing, separate reviews of CBT and pharmacotherapy.	We changed SSRI. We disagree with reviewer that the report has significant conceptual errors. Issues suggested by them (subgroup analyses, coding, looking at secondary publications) have all been already done when feasible. The fact that the conclusions of the report are similar to other reviews is not a problem of the report.
Peer Reviewer # 4	Clarity and Usability	I think a great deal of work needs to be done to make this really available to the public.	Thank you. The final report will be posted online.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #1	General	This paper clearly represents a staggering amount of work and the authors should be commended for their attempt to synthesize such a large and varied literature. In part because of the enormous effort that obviously went into this project, it seems important to get the paper as useful as it can be to people who need this information. Here, in my view, there are some problems. This review is not a particularly user-friendly given the number of acronyms, references, and tables. The writing style is also lacks some cohesion. The review also restricts itself only to medications and traditional psychotherapy despite there being evidence in some other modalities such as mindfulness, physical activity, and nutrition. It could be very useful to at least mention these other areas and remind clinicians that there are more potential avenues for intervention. Too often, the manuscript reads as a dry list of the evidence that ends up equating studies that don't seem in the same ballpark. For example, there is far more evidence for sertraline than paroxetine in treating child anxiety disorders but the review treats the two as essentially identical. There are a fair number of typos and formatting inconsistencies (sometime writing out “six” versus using the number “6,” capitalizing names of drugs occasionally but not always) that need to be addressed. There are also some odd expressions like “rate of diagnosis free (pg 27, line 28)” that will seem strange to many readers.	We thank the reviewer’s comments. We extensively edited the current report and updated results. Due to the limited time and reporting requirements, we are not able to dramatically change the structure of the report. However, we willreorganize/rewrite the report for journal publications. With the new categorization of intervention we included the evidence for non-CBT therapies. We examined our search criteria and updated search till February 1, 2017. We added 6 more studies in the revision.

Commentator & Affiliation	Section	Comment	Response
TEP 1	General	The largest, and a quite significant, concern that I have is with the way that psychotherapy treatments are being described and grouped for analyses. Specifically, several of the treatments listed as "non-CBT"— (e.g., PCIT is certainly a form of CBT, as are many PST programs since problem solving is a key component of CBT interventions; single-session therapy is CBT; and modified CBT for Autism would still be CBT). Further examining of the studies detailed in Table E.4 further show this confounding as several of the studies are listed as "other therapy" but then the actual components of the program are cognitive and behavioral (e.g., exposure). All of this has major implications for the entire analytic approach and resulting interpretation of these domains and the comparisons among them. For example, as written the lack of differences between CBT and Other Psychotherapy may just be because you predominantly comparing CBT to CBT at present. These studies should be reviewed by experts in the field who can reliably assign them to categories and analyses should be recalculated.	In this revision, we checked all of the included studies and re-evaluated/re-categorized all of the interventions and outcomes by two clinical experts. We hope this revision has addressed all of your concerns.
TEP 1	General	Also, the key questions that occur at the beginning of the manuscript cover several domains that do not actually end up being examined in the report (e.g., the different types of comorbidity). It would be clearer if the authors specified what is in the report only (or perhaps noted the other domains as aspirational but not feasible to examine).	We thank the reviewer's comments. The key questions were listed in the protocol and were used to guide the whole process. We summarized the findings in abstracts. Hopefully this would address your concerns.
Peer Reviewer # 2	General	This report is clinically meaningful. Target population and audience and key questions are all explicated clearly and in an up-front manner.	Thank you.
TEP 2	General	Target population and audience are explicitly identified and key questions are appropriate and explicitly stated. The clinical utility of the report is limited by the lack of data for use in more meaningful subgroup analyses. The authors note this in their Limitations section.	Thank you.

Commentator & Affiliation	Section	Comment	Response
TEP 2	General	Thank you for the opportunity to review the draft of the comparative effectiveness review about Anxiety in Children. In general, I found the review thorough and an accurate reflection of the current evidence base in this field. I have one question/comment about the analytic procedures for part of the review, one caveat that I think needs to be applied to one of the results, and the rest of my comments are wording edits/suggestions. Regarding the analytic procedures for the subgroup analyses, the contrasts that were tested did not conform to what I would have expected from the perspective of trying to guide the field about which intervention to choose for which patients. Specifically, I was surprised to see that patients with various types of comorbidity (i.e., ADHD, autism, depression) were being compared to each other and there were no contrasts that addressed the question of relative effectiveness for cases with and without each type of comorbidity. For example, if I am a provider trying to choose among interventions for a child with comorbid ADHD, I do not really care how treatment X performs for that subgroup relative to patients with another form of comorbidity that my patient does not have (e.g., autism). Is it a limitation of this type of review and the data available to you that these more clinically relevant comparisons cannot be done? Even if the type of subgroup analysis that I am requesting cannot be done, I do not think that the results from the comorbidity subgroup analysis should be highlighted in the abstract because they feel esoteric.	We thank the reviewer for the nice comments. In terms of the suggestion about subgroup analyses based on comorbidities, we conducted more subgroup analyses in the revised report, including patients with comorbidity (vs. no comorbidity), ADHD, autism, school refusal. We also conducted more subgroup analyses in Page 22-23.

Commentator & Affiliation	Section	Comment	Response
TEP 2	General	Wording edits: (Note: the page numbers I reference here are the page numbers in the original manuscript document, NOT the page numbers at the top left) 1.p. viii, line 17: insert the word “may” so that the phrase reads: “…anxiety symptoms and improve functioning but may result in adverse events…” 2.p. 1, line 10: use of lifetime prevalence twice in the same sentence is confusing. Consider rephrasing the first prevalence estimate for ages 13-18 since that timeframe is not truly “lifetime.” 3.P. 1, line 25: replace word “selective” with “serotonin” (this issue comes up several more times throughout the manuscript) 4. P. 2 Table 1: The classification of therapies into CBT and non-CBT does not conform in all cases to how I would have classified therapies. In particular, motivational interviewing and ACT-based treatments (many of which are considered “Third Wave”) I think of as still being in the broad category of CBT. Also, later on p. 22 line 50 when “modifications of CBT for patients with autism spectrum disorder” is being grouped as a “non CBT” treatment does not make sense to me. Maybe the whole non-CBT category needs another name? 5.P.2, Table 1: Name for the 1st drug class listed needs the same edit as mentioned in item 3 above. 6. P.6, line 12: the end point age of 18 listed here does not agree with ones listed on p. 44 (line 9) and p. 45 (line38) that say 17. 7.P. 7, line 29: missing “à” before the word priori 8. P. 8, line 28: word missing? Phrase now reads: “…we rated SOE by one level.” Would “rated down” or “lowered” instead of “rated” be accurate to what the authors meant? 9. P.15, line 41 in second bullet point: replace “compared” with “more than” 10. P.15, line 47: same edit at #9 above	Thank you. We changed these.

Commentator & Affiliation	Section	Comment	Response
TEP 3	General	The report is well done. It clearly stated its objectives, target audience, and methods. One comment on the abstract. I wondered about the need to mention the sub-analyses since I believe it was exploratory. The sentence about CBT impact on anxiety and ADHD versus Autism was confusing and seemed to suggest it was also studied in schools. But I realize the comparison was about delivery in different settings. In the conclusion of the abstract I might consider stating as a class the SRIs have evidence of effectiveness. The mechanism of these compounds involve serotonin but really don't know how they work.	Thank you for your comments and suggestions. We revised the sentence in the abstract.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer # 3	General	Anxiety disorders are the most prevalent and most undertreated pediatric mental health problem. Understanding the comparative efficacy and harms of psychosocial and pharmacological interventions has important public health implications. However, while the overall focus on the report is significant and meaningful some of the more ambitious secondary questions -- such as subgroup analyses -- are not realized in this analysis. As such the overall impact of the report is diminished and the content is much more similar to other existing evidence-based reviews of the literature. Furthermore, omission of methodological details (and apparent errors) reduce confidence in the quality and final conclusions of this review relative to others. In brief, our primary overarching concerns relate to lack of clarity and consistency within and between the methods, results, discussion, and tables, as well as what we believe are serious weaknesses in category definitions of interventions. Given the interplay between issues raised across the various sections of the review, our comments will be somewhat repetitive in the sections that follow, with the bulk of our comments inserted into the discussion of the methods section.	We thank the reviewers' comments and suggestions. In current report, we made extensive changes. We hope these would have addressed your concerns.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer # 4	General	Not clear who the target audience is give the methods and structure of the paper. I think the methodology is so arcane. The extraordinary efforts to be comprehensive results in a paper that is very difficult to read and understand. There are so many small studies that are included and the large (by the field's standards) are ignored. The key questions are fine. The recommendations lack sophistication of methodological limitations if recommended studies such as long term studies or studies of moderators or mechanistic studies. There are a number of typos and mismatch of refs in the body and those in the ref list The disorders included were a problem - too broad	It is true that the focus on comprehensiveness can make evidence reports long and difficult to read. We also conducted comprehensive search of the literature and are confident most, if not all, relevant studies were included . We corrected typos and cited correct references. We also substantially revised the report, especially the discussion section.