Dear Dr. Loder and the BMJ Editorial Team,

We thank you for your thoughtful review of our manuscript and are pleased to resubmit it after revisions. We agreed with the vast majority of comments and made several changes to the paper in response.

Below, we address each reviewer/editor comment (comment left intact; our responses are bulleted):

Reviewer: 1

Thank you for the opportunity to review this interesting paper. It shows that the existing literature demonstrates similar efficacy between CBT and SGAs. The study design and conduct are appropriate and the paper is well written. It has some overlap with Spielmans et al., 2011 (ref#25) in that it includes many of the same studies although, as the authors point out, it differs in its lack of emphasis on "bona-fide" vs. "non-bona fide" psychotherapies and also in it's a priori attempt to examine outcomes beyond response and remission (adverse effects, relapse, quality of life etc.). The authors are correct to highlight in the introduction that primary care physicians need better evidence to select appropriate treatments for patients. This study is an important contribution to the literature and deserves to be published.

• We thank Dr. Sinyor for his thoughtful review and recommendation. We address his specific comments below.

Some specific comments:

1) A key finding of the study is the low overall numbers/power. The authors mention this but it may be worth greater emphasis. For example (page 10, line 15), remission rates to CBT vs. SGAs were 47.9% vs. 40.7%. While this was not a statistically significant result due to low numbers, a true 7% difference may be clinically significant. I would favour a stronger statement of caution in interpreting the results such as amending the conclusion in the abstract to say that while benefits were similar between the two treatments, low numbers precluded the analysis from detecting small but potentially meaningful differences.

• We agree that a nonsignificant (statistically speaking) finding may still have meaningful clinical relevance, particularly when results are considered in a population-level context from a public health perspective. We have revised the abstract to emphasize this point.

2) Likewise the authors could consider a greater emphasis on the fact that the strength of evidence (Table 3) is quite low overall.

• As above, we have added language to the abstract to encourage caution in interpreting the results given the low SOE.

3) In the first paragraph of the introduction, the authors descirbe APA guidelines (reference #3) which define psychotherapy as at least 8 visits of an average of 30 minutes. All of the CBT therapists with whom I am familiar have sessions lasting 45-60 minutes and a reader would likely assume that this was the CBT examined here (likewise, 8 sessions is probably "underdosed" although all of the included studies had at least 14 which is reassuring). I have not read all of the studies included in the review in detail, but I would be surprised if any of them allowed more abbreviated sessions. If so, it would be important to highlight that and consider running a separate analysis excluding them.

• We appreciate this point and agree that results from studies using sessions shorter than 30 minutes may differ from those using more traditional durations. We reviewed the included studies' session durations and found that seven (64%) provided sessions at least 50 minutes long; two (18%) did not report the CBT session duration; and two (18%) offered 30 to 40-minute sessions. The two studies that used shorter sessions (Lam et al., 2013 and Mynors-Wallis et al., 2000) reported no statistically significant differences in response, remission, or reduction of symptom score. Numerically, the results very slightly favored SGA in most cases, which supports the reviewer's implication of "underdosing" via less than standard session duration. We ran sensitivity analyses that removed those studies and found no significant (statistically or clinically) difference in results for response or remission:
o Remission: changed from 0.98 (0.73 to 1.32) to 0.87 (0.64 to 1.32)
o Response: changed from 0.91 (0.77 to 1.07) to 0.91 (0.77 to 1.08)
Removing those studies from the analysis of change in HAM-D did not change the direction or statistical significance of the results, but the effect size of CBT increased from -0.38 to -1.70 (though the resulting analysis contained only 2 arms from a single study).
As a result of the above, we feel comfortable with the results and conclusions from our original analyses.

4) The authors could consider elaborating on page 12 line 32 in which they note their results are consistent with APA/VA guidelines. They would also support, for example, the UK NICE guidelines which suggest that CBT may even be a preferred first option in mild to moderate depression

• We agree, and we have added the mention of the NICE guidelines.

5) The authors variously describe CBT/SGA as "initial treatment" "an initial treatment attempt" and "initial outpatient treatment". I assume this means the initial treatment of a new major depressive episode (as opposed to augmentation or subsequent trials following a treatment failure). It could be interpreted as a patient's first treatment for MDD which I don't believe the authors intend (i.e. I assume subjects were allowed to be included if they had a history of prior MDEs that had remitted in the past). Please clarify in the methods.

• We appreciate this request for clarification, and we have edited the text to indicate that the patients were undergoing an initial treatment for the current (not necessarily first-ever) MDE.

6) The authors' conclusion that, given similar efficacy, the decision to choose between treatments should include

patient preference, cost, availability and expectancy effects is reasonable. They should also consider referencing the following study: "McGrath CL, Kelley ME, Holtzheimer PE, Dunlop BW, Craighead WE, Franco AR, Craddock RC, Mayberg HS. Toward a neuroimaging treatment selection biomarker for major depressive disorder. JAMA Psychiatry. 2013 Aug;70(8):821-9." These results are very preliminary, but if a biomarker for response to CBT vs. SGAs were found, it would be a crucial tool for directing treatment.

• We have added this point.

Reviewer: 2

Comments:
Dear authors and dear editor,

This paper is about a crucial topic. Although being not very innovative, this study complies with the best standards of meta-analysis. I found no major issue as regard the conduct of the systematic review and the meta-analysis. My minor methodological comments are listed below.

• We thank Dr. Naudet for the thorough review and insightful comments, which we address below.

Do authors have check the following database : http://www.psychotherapyrcts.org/index.php?id=25. It is a very helpful platform for meta-analyses in the field of psychotherapeutice research?

• This is an excellent compilation of the published research, and we thank Dr. Naudet for bringing it to our attention. Of the 18 records with relevant comparisons, we have included eight. Of the remaining 10, we excluded four as ineligible for our review. In response, we have reviewed the remaining six articles and found that none of them meet our criteria due to wrong condition (i.e., not solely MDD).

When authors states “Evidence indicates that all SGA have the same efficacy”, I agree but it is not so simple. This result has not been replicated by Cipriani et al. (Lancet) and there is a considerable debate in the litterature. Please be less affirmative and nuance your opinion. I repeat that I agree with your opinion but it is just an opinion. In the other hand a previous study has stated that there all psychotherapies have the same efficacy (Plos Medicine, Barth et al.). One can regret that the other psychotherapies were not included in this study. In fact network meta-analyses allow adding an individual effect for a given treatment and a class effect. This would have been useful for both psychotherapies and for antidepressive agent. But I acknowledge that these important questions were not objectives for the present paper.

• We have edited the text regarding the efficacy across all SGAs; it is now less definitively stated.
• Although this paper only examined CBT, the full report from which it is derived has several other psychotherapies. We have added clarification that the full report also contains comparisons between SGAs and psychotherapies other than CBTs.

In the PRISMA chart, please detail the reasons of exclusion between the trials include for the full evidence synthesis (n=44) and the trials included in the present meta-analysis (n=11). It would be very informative.

• The 33 studies not included in the present paper compare SGAs with non-CBT psychotherapies and complementary/alternative therapies (e.g., acupuncture, St. John’s wort, exercise). We have updated the figure with this explanation.

But here is the most important point. Would the antidepressive agents have been different to placebo in these populations? Indeed, baseline scores between 16 and 23 of baseline score on a Hamilton scale are not very high and the Kirsch meta-analysis and the Fournier meta-analysis suggested that there were differences between antidepressants and placebo only for severe (or very severe) MDD. If I also agree that these data are disputed in the literature, the absence of a placebo control is a major limitation for this study, especially when one concludes to non-inferiority between interventions. Is it a non inferiority because both are effective or because both are not effective? This is a crucial in terms of interpretation and in terms of clinical relevance. It is really a problematic point.

• Although we appreciate the reviewer’s point, we did not set out to determine whether either treatment is better than placebo; rather, we undertook the comparison of “real-world” interventions for patients with MDD. While we concluded that CBT and SGAs are equally effective, we do agree that CBT and SGAs could also be equally ineffective. We have added a statement to this effect to the Discussion.

At last the question of publication bias is quite hard to address with such a small number of studies. Nonetheless, it could have a major impact on the observed results of the meta-analysis.

• As written in the methods section (both original and revised versions), we did examine funnel plots of the meta-analyses for evidence suggesting publication bias, and we found none; we also noted Dr. Naudet’s point that publication bias is difficult to assess with few included studies. Without reason to suspect significant publication bias, we do not feel that speculating on the potential effects is informative.

Allegiance bias is a very important bias in psychotherapeutic research. How authors have tried to explore this bias. It would be interesting to add allegiance in the description of studies.

• We agree that allegiance could be an interesting concept to explore in terms of potential bias, but we were unable to find any descriptions of allegiance in the included studies.

At last, one of the major results of the study is not the absence of evidence but is the quality (bad) and the number of studies (small) on such a crucial question. This should be emphasised in the discussion.

• We agree and have edited the text to encourage more caution in interpreting the results given these limitations.

To summarize my point of view, I have just a few concerns concerning this meta-analysis that must be addressed in the discussion to allow for a more reasonable interpretation of the findings.

On the other hand, many studies were conducted on the same topic and I'm not sure that the data presented here add a lot to the question. But it rather an editorial question than a scientific one. Here are the following studies that (more or less) reach the same conclusions, except for combined treatment:

1/ Cuijpers P, Sijbrandij M, Koole SL, Andersson G, Beekman AT, Reynolds III CF (2013). The efficacy of psychotherapy and pharmacotherapy in treating depressive and anxiety disorders: a meta-analysis of direct comparisons. World Psychiatry, 12, 137–148.

-> "Althougth more large sample and older antidepressants, the overall effect size indicating the difference between psychotherapy and pharmacotherapy after treatment in all disorders was g=0.02 (95% CI: -0.07 to 0.10), which was not statistically significant."

2/ Cuijpers P, Berking M, Andersson G, Quigley L, Kleiboer A, Dobson KS (2013). A meta-analysis of cognitive behavior therapy for adult depression, alone and in comparison to other treatments. Canadian Journal of Psychiatry, 58, 376-385.

-> "We did not find any indication that CBT was more or less effective than other psychotherapies or pharmacotherapy. Combined treatment was significantly more effective than pharmacotherapy alone (g = 0.49)."

3/ Spielmans GI1, Berman MI, Usitalo AN. Psychotherapy versus second-generation antidepressants in the treatment of depression: a meta-analysis. J Nerv Ment Dis. 2011 Mar;199(3):142-9. doi: 10.1097/NMD.0b013e31820caefb.

"Bona fide psychotherapy appears as effective as SGAs in the short-term treatment of depression"

4/ Cuijpers P, Sijbrandij M, Koole SL, Andersson G, Beekman AT, Reynolds III CF (2013). Adding psychotherapy to antidepressant medication in depression and anxiety disorders: A meta-analysis. World Psychiatry, 13, 56-67.

"We conclude that combined treatment appears to be more effective than treatment with antidepressant medication alone in major depression, panic disorder, and OCD."

• We appreciate the chance to clarify what our works adds to the comprehensive collection and synthesis of studies noted above (the publications each draw from the same, growing collection of trials). Specifically, our eligibility criteria (and, therefore, our sample of studies) differed in the following ways, which reflect our interest in primary care-relevant information:
o Our sample includes only studies where the comparison arm involves second generation antidepressants, the predominant choice for primary care physicians (whereas in the Cuijpers World Psychiatry analysis, for example, 13/32 studies used only tricyclic antidepressants or monamine oxidase inhibitors);
o Our inclusion criteria required the use of evidence-based psychotherapies (so, for example, did not include non-directive supportive counselling, which were included in the Cuijpers studies);
o Our selection criteria required that the patients did not already have a treatment failure in the current episode, reflective of patients seen in primary care who get a initial treatment attempt from the primary care physician;
o The above studies did not include data on harms, which we report here.
o Our review included articles published through January 2015, which updated the above by adding 3 new relevant articles.
• This information was addressed in the first paragraph of the discussion. To clarify the above, we have added to the description of what our analyses add in that same paragraph.

Please excuse my english but I tried to write this review as rapidly as possible. Please also excuse the delay in answering due to an important clinical activity.

All the Best

Additional Questions:

BMJ EDITORIAL COMMENTS

* We thought the concern about lack of placebo comparison might be less relevant in practice as its unlikely that doing nothing would be an option in major depressive disorder, presumably because there is existing evidence of benefit. You might consider making that point in the discussion.

• We agree and have done so.

* We would like you to explain in more detail how your work adds to other recently published papers that seem to address a very similar research question, especially since it seems that US guidance is already consistent with your findings. For example, Cuijpers P, Sijbrandij M, Koole SL, Andersson G, Beekman AT, Reynolds III CF (2013). The efficacy of psychotherapy and pharmacotherapy in treating depressive and anxiety disorders: a meta-analysis of direct comparisons. World Psychiatry, 12, 137–148. Cuijpers P, Berking M, Andersson G, Quigley L, Kleiboer A, Dobson KS (2013). A meta-analysis of cognitive behavior therapy for adult depression, alone and in comparison to other treatments. Canadian Journal of Psychiatry, 58, 376-385.
Spielmans GI1, Berman MI, Usitalo AN. Psychotherapy versus second-generation antidepressants in the treatment of depression: a meta-analysis. J Nerv Ment Dis. 2011 Mar;199(3):142-9. doi: 10.1097/NMD.0b013e31820caefb.

Cuijpers P, Sijbrandij M, Koole SL, Andersson G, Beekman AT, Reynolds III CF (2013). Adding psychotherapy to antidepressant medication in depression and anxiety disorders: A meta-analysis. World Psychiatry, 13, 56-67.

• Please see our response to reviewer #2's very similar comment above.

* The quality of the evidence seems to remain low and some of these confidence intervals are wide. Just over 1 thousand patients have been randomised so far. So despite all of the meta-analyses it seems the question is not entirely settled. Perhaps this matter deserves comment, either by you in the discussion or by an editorialist. To be provocative, why is the government funding a meta-analysis when previous work might suggest that it's a definitive trial that is needed?

• One of the key drivers of the report was the need to synthesize existing evidence that is most relevant to primary care. Primary care physicians provide the largest number of antidepressant prescriptions and account for most of the near doubling in the use of antidepressants over the past decade.30 Accordingly, much of this treatment may be for patients with either threshold or mild MDD, suggesting a risk of overtreatment for this group. At the same time, primary care physicians appreciate that other potentially effective interventions are available. According to the topic nominators, primary care physicians require an evidence base identifying the comparative effectiveness of the available treatments for depression to increase the likelihood that treatments are selected and managed correctly.
• Although we addressed that point toward the end of the Introduction, we appreciate that it warrants confirmation in the Discussion. We hope our edits in response to this and similar comments about what this work adds to the body of evidence do so satisfactorily.
• That written, we certainly welcome comment from an editorialist about the issue.

* Our statistician thought that the results could be laid out better and more information should be given rather than just a reference to the full report (which may have much of the information required to properly interpret the findings). Some of her points re the study design and analysis are given below:

1. Why were only studies with a minimum sample size of 500 included in the harms analysis?

• There is good evidence that RCTs cover common adverse events adequately, and we did not feel the need to replicate such findings by pooling many small observational studies. However, detection of rare but serious adverse events (i.e., events for which the comparative rates of harms may remain unknown) requires large observational studies.

2. Why was the 'worst-case'assumption only used for studies with attrition? Did the authors assume 'worst-case' for losses in each arm? Why not also have 'best-case' and also the combinations that best favour each arm (ie. all attrition have favourable outcome in one arm and unfavourable in the other)? The numbers of losses in each study and arm should be given. (The discussion states that several studies have very high attrition but the actual numbers are not apparent.)

• We imputed data only for studies in which the number analyzed was less than the number randomized in order to minimize bias due to individual studies' rates of attrition and method of handling missing data.
• We chose not to perform a "best-case" analysis because we wanted to take a cautious, conservative approach; we did not think that a "best-case" assumption or one of favorability in one outcome and not the other would be particularly useful to inform clinical decisions.

3. The authors use the Cochrane risk of bias tool but give only an overall risk of low, medium or high. Details of the subscores and methods of combining to a single assessment should be given (these may be in the full report (reference 26) but I think details should be given here too.

• We have added the full risk of bias assessments to the online supplement. (They are, indeed, also in the full report.)

4. One included study (ref 32) had 2 comparison arms for CBT (Ct and REBT) and the same comparison SGA group. How is the correspondence between the estimates from these 2 arms (same control sample and within the same study) accounted for in the random effect modelling?

• In order to avoid double counting of the patients in the SGA arm, we split the SGA sample in half and compared each half with a treatment arm. For dichotomous outcomes, we split the number of patients who responded/remitterd in half; for the continuous outcome, we assumed the same mean for both halves as was reported for the overall group.

5. The emphasis on interpretation should be the combined statistics and confidence limits for the RR obtained from the meta analysis.

• We have added mention of the analysis results to the early part of the Discussion.

6. Table 3 would benefit from having the RR (ci) included for each comparison

• We have added the effect sizes (where estimable) to Table 3.

7. A network analysis may be more informative, was this considered?

• We did consider the use of network meta-analysis. Because we found randomized, head-to-head comparisons, even though they are few, we feel that pairwise meta-analysis is more appropriate. Network meta-analysis requires some key assumptions, most of which we are unable to verify.

8. The study does not show that there are similar beneficial effects (abstract conclusions), no significant difference does not imply equivalence nor that both are better than nothing. Please reword your conclusion to better reflect this and mention this in the abstract as well.

• We have reworded the conclusion and edited the abstract to address this point.

End of responses

We hope that you find our responses and the corresponding revisions satisfactory; we look forward to your reply!
Halle Amick, on behalf of the EPC team

1. BMJ response.docx PDF HTML