PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	The efficacy of dapagliflozin combined with hypoglycemic drugs in treating type 2 diabetes mellitus: meta-analysis of randomized controlled trials
AUTHORS	Sun, Yu-nan; Zhou, Yi; Chen, Xi; Che, Weng-si; Leung, Siu-wai

VERSION 1 - REVIEW

REVIEWER	Peter Watson Medical Research Council, UK.
REVIEW RETURNED	10-Jan-2014

GENERAL COMMENTS	There are a couple of instances where the grammar could be improved (see comments to authors). I'd mention in the abstract that HbA1c and FPG both measure sugar levels. It would also be very helpful to describe further via a funnel plot and a SMD confidence interval the results of the trim and fill adjustment for publication bias described at the top of page 9 and in the third paragraph on page 10. It is also customary and a good idea to add in 95% confidence intervals to funnel plots such as those in Figure 4 (sloping lines at each side) to further assess the symmetry of the study effect sizes. Looking at Figures 3a-3c one can see that there is such a consistent negative signage of effects in the 95% confidence intervals as to suggest grounds for a robust effect of DPA which must be regarded as a strength in this paper. Since the authors then address publication bias the conclusion of improved glucose control/body mass seems justified although a few further details of the results of the trim and fill adjustment would be welcome The efficacy of dapagliflozin combined with hypoglycemic drugs in treating type 2 diabetes mellitus: meta-analysis of randomized controlled trials. bmjopen-2013-004619 The results in Figures 3a-3c show consistency with nearly all the studies having 95% confidence intervals for standardised mean differences containing only values less than zero. This in itself suggests Dapagliflozin is beneficial as an add-on drug for a variety of outcomes and time periods given that there is no publication bias or, in the one case where this is detected, using the trim and fill adjustment (page 10, third paragraph). Page 5. It would be helpful to define in the abstract on page 2 that

	HbA1c and FPG measure glucose levels. Page 6, start of data synthesis paragraph. Given that the freeware, R, was used for performing other statistical tests and contains a procedure, metafor (Viechtbauer (2007)) for doing meta-analyses, I was surprised to see other non-R software used for doing the meta-analysis. Page 6. (Line 4 of the data synthesis paragraph) It is implied that random effect models were used for computing overall effect sizes regardless of the amount of heterogeneity between the studies. Is this the case and if so what is the usefulness of the I^2 statistic measure of between study heterogeneity? For example, one might have used the I^2 measures to determine which effect estimate was used using fixed effect estimates for small amounts of study heterogeneity (with low I^2) and random effect sizes for larger effects of heterogeneity (with large I^2). Page 6 (last sentence of data synthesis paragraph). One could add in the reference (Higgins et al. (2003)) motivating the cut-offs used for I^2. Page 7. The last sentence of the second full paragraph is grammatically incorrect. Page 8 (fourth last line of first full paragraph). Could delete the second mention of 'two durations' from the sentence 'The two duration subgroups on two durations…'. Page 9. First full paragraph. The meta-regression whose results are also given in Table 3 (page 20) is not described. What is it pooling and across which studies? Is it pooling body mass correlations from times <=24 weeks and >24 weeks across studies? In the last sentence in this first full paragraph R-squared is given as equal to 1. Is this correct and, if so, is such a strong relationship plausible or does it suggest a limitation in the data? Such a R-square would imply that body weight differences at <= 24 weeks tell you exactly the body weight difference >24 weeks. Is the 'efficacy was time dependent' (line 5 of first full paragraph) simply saying the body mass differences between the groups declined with time to the point where the DPA is being regarded as having a decreasing influence over time. If so, is this a tenable conclusion given we are comparing just two time points (<= 24 weeks and > 24 weeks)? Page 10. (third paragraph) Where are the trim and fill results mentioned here reported? These results should be reported in the results section on page 9 as they adjust for the publication bias reported in lines 4 and 5 at the top of page 9. The trim and fill procedure is presumably that of Duval and Tweedie? It would be helpful in the analysis paragraph on page 6 to mention that a trim and fill procedure is used if needed to correct for publication bias. One might also mention on page 6 (for those unfamiliar with meta-analyses) what we mean by publication bias ie the danger that only studies with larger effects are published. Could you elaborate on the sentence at the end of this third paragraph on page 10 concerning the meta-regression finding DPA having significant time-dependent effects. Is this saying DPA levels

	change over time (presumably higher levels <= 24 weeks compared to >24 weeks) irrespective of FPG and body weight. If so is this suggesting the reason why the 'efficacy is time-dependent' on page 9 (first full paragraph) namely that since DPA levels decrease over time its effect decreases thus reducing differences in HbA1c, FPG and body weight between the DPA and control groups. Does this suggest a timeframe suggesting how long one should wait before topping up levels of DPA with further injections to control glucose levels? Page 26. Is it possible to add in the 95% confidence limits to the funnel plots in Figure 4 to give a stronger indication of publication bias? I would also add in the adjusted funnel plot with extra studies denoted as added by the trim and fill procedure described in the third paragraph on page 10. Pages 23, 24, 25, 27 and 28. STANDARDISED means difference would be more informative than 'mean difference' to describe the effects stated in the tables. In passing I would draw attention to Peters et al. (2010) who criticise the use of Egger's test which is used here as a measure of publication bias. References Higgins, J.P., Thompson, S.G., Deeks, J.J and Altman D.G. (2003). Measuring inconsistency in meta-analyses. BMJ 327 557-560. Peters, J. L., Sutton, A. J., Jones, D. R. and Abrams K. R. (2010). Assessing publication bias in meta-analyses in the presence of between-study heterogeneity. Journal of the Royal Statistical Society A 173(3) 575-591. Viechtbauer, W. (2007). Confidence intervals for the amount of heterogeneity in meta-analysis. Statistics in Medicine 26(1) 37-52.

REVIEWER	Abd A Tahrani Clinician Scientist in Diabetic Medicine, University of Birmingham, UK
REVIEW RETURNED	08-Feb-2014

GENERAL COMMENTS	The manuscript is well written. The methodology is strong and appropriate. Major Comments: In the strengths and limitations the authors state "This study is the first meta-analysis to focus on the weight gain issue of dapagliflozin". What does this meta-analysis add to those published previously such as Goring et al. Diabetes Obes Metab. 2013 Nov 14. doi: 10.1111/dom.12239? The meta-analysis included a mixture of RCTs some of them were the same study reported at different time points, why the authors didn't just included the final analysis with the longest follow up as these likely to be more meaningful clinically. Minor comments In key messages, the authors state "Dapagliflozin as an add-on drug improves the control of HbA1c and FPG levels in type 2 diabetes mellitus (T2DM) patients with little risk of weight gain"; overall I agree but I do not think the results support the statement regarding

	little weight gain. Teh results suggest that Dapa treatment was associated with weight loss. In order to justify the authors statement they need to show us the risk of weight gain in Dapa treated patients which is not included in the current manuscript. The issue of publication bias in regards to weight loss but not in regards to HbA1c in interesting. I agree that the publication bias is evident based on the funnel plot. Although this is significant statistically I am not sure whether a true bias exists. All the Dapa RCTs reported the change in glycaemic measures (HbA1c) and weight; in addition the primary outcome is usually change in HbA1c. I am not aware that any studies were conducted with weight being the primary outcome; hence it is unlikely for studies not to be published because of lack of effect on weight.

VERSION 1 – AUTHOR RESPONSE

REVIEWER 1:

“Please state any competing interests or state ‘None declared’: None declared.”

>> Thanks for your advice. The statement of competing interests has been added.

“There are a couple of instances where the grammar could be improved (see comments to authors).”

>> Thanks for your correction. The grammar has been improved according to your comments.

“I'd mention in the abstract that HbA1c and FPG both measure sugar levels.”

>> According to your suggestion, the main purpose of both HbA1c and FPG to measure the glycaemic level has been added in the Methods of ABSTRACT.

“It would also be very helpful to describe further via a funnel plot and a SMD confidence interval the results of the trim and fill adjustment for publication bias described at the top of page 9 and in the third paragraph on page 10.”

>> According to your suggestion, the funnel plot has been updated in Figure 6. An estimated mean difference and 95% confidence intervals of the trim-and-fill adjustment on body weight have been also added to the Risk of bias across studies of RESULTS.

“It is also customary and a good idea to add in 95% confidence intervals to funnel plots such as those in Figure 4 (sloping lines at each side) to further assess the symmetry of the study effect sizes.”

>> According to your suggestion, 95% confidence intervals have been added to funnel plots in Figure 6.

“Looking at Figures 3a-3c one can see that there is such a consistent negative signage of effects in the 95% confidence intervals as to suggest grounds for a robust effect of DPA which must be regarded as a strength in this paper. Since the authors then address publication bias the conclusion of improved glucose control/body mass seems justified although a few further details of the results of the trim and fill adjustment would be welcome (see further comments below).”

>> According to your suggestion, we have added the results (including the estimated mean difference and 95% confidence intervals) of the trim-and-fill adjustment in the Risk of bias across studies of

RESULTS.

“The results in Figures 3a-3c show consistency with nearly all the studies having 95% confidence intervals for standardised mean differences containing only values less than zero. This in itself suggests Dapagliflozin is beneficial as an add-on drug for a variety of outcomes and time periods given that there is no publication bias or, in the one case where this is detected, using the trim and fill adjustment (page 10, third paragraph).”

>> In accordance with your suggestion, the use of the trim-and-fill was described in the third sentence from the last in the Data synthesis and analysis in METHODS. We used the trim-and-fill adjustment on body weight, where publication bias was detected. The detail explanation was in the Risk of bias across studies of RESULTS.

“Page 5. It would be helpful to define in the abstract on page 2 that HbA1c and FPG measure glucose levels.”

>> Yes. The purpose of HbA1c and FPG to measure glycaemic level has been added in the Methods of ABSTRACT.

“Page 6, start of data synthesis paragraph. Given that the freeware, R, was used for performing other statistical tests and contains a procedure, metafor (Viechtbauer (2007)) for doing meta-analyses, I was surprised to see other non-R software used for doing the meta-analysis.”

>> Thanks for your advice. We did use both Review Manager and R’s metafor package to conduct meta-analysis although the plots generated by the Review Manager were presented because they look more informative and clear in drawing. Based on your advice, we have revised the description in the Data synthesis and analysis of METHODS.

“Page 6. (Line 4 of the data synthesis paragraph) It is implied that random effect models were used for computing overall effect sizes regardless of the amount of heterogeneity between the studies. Is this the case and if so what is the usefulness of the I^2 statistic measure of between study heterogeneity? For example, one might have used the I^2 measures to determine which effect estimate was used using fixed effect estimates for small amounts of study heterogeneity (with low I^2) and random effect sizes for larger effects of heterogeneity (with large I^2).”

>> We used I^2 to measure the heterogeneity level, but did not use it as a criterion to choose random-effect model. According to some meta-analysis texts and papers (e.g. Schmidt et al., 2009), the choice of model between the fixed-effect and random-effects model should be made in advance according to our expectation of heterogeneity rather than the results of I^2. In this study, we chose the random-effects model also due to the model properties to be more conservative and generalizable.

“Page 6 (last sentence of data synthesis paragraph). One could add in the reference (Higgins et al. (2003)) motivating the cut-offs used for I^2.”

>> Thanks for your suggestion. The reference has been added accordingly (reference No.19).

“Page 7. The last sentence of the second full paragraph is grammatically incorrect.”

>> Thanks for your correction. The last sentence in the Study characteristics of RESULTS has been

revised.

“Page 8 (fourth last line of first full paragraph). Could delete the second mention of ‘two durations’ from the sentence ‘The two duration subgroups on two durations…’.”

>> Thanks for your suggestion. The second mention of ‘two durations’ has been deleted.

“Page 9. First full paragraph. The meta-regression whose results are also given in Table 3 (page 20) is not described. What is it pooling and across which studies? Is it pooling body mass correlations from times <=24 weeks and >24 weeks across studies? In the last sentence in this first full paragraph R-squared is given as equal to 1. Is this correct and, if so, is such a strong relationship plausible or does it suggest a limitation in the data? Such a R-square would imply that body weight differences at <= 24 weeks tell you exactly the body weight difference >24 weeks.
Is the ‘efficacy was time dependent’ (line 5 of first full paragraph) simply saying the body mass differences between the groups declined with time to the point where the DPA is being regarded as having a decreasing influence over time. If so, is this a tenable conclusion given we are comparing just two time points (<= 24 weeks and > 24 weeks)?”

>> Thanks for your advice. The meta-regression was conducted on different follow-up durations including 12th, 24th, 48th, 102nd, 104th weeks, which has been explained as you advised with all details in the Data synthesis and analysis of METHODS. In the Synthesis of results from individual studies of RESULTS, the results of meta-regression were described in the last two sentences of the subsections of HbA1c, FPG, and Body weight. The R-square may be overestimated. In addition, we mentioned this as a limitation in DISCUSSION.

“Page 10. (third paragraph) Where are the trim and fill results mentioned here reported? These results should be reported in the results section on page 9 as they adjust for the publication bias reported in lines 4 and 5 at the top of page 9. The trim and fill procedure is presumably that of Duval and Tweedie? It would be helpful in the analysis paragraph on page 6 to mention that a trim and fill procedure is used if needed to correct for publication bias. One might also mention on page 6 (for those unfamiliar with meta-analyses) what we mean by publication bias ie the danger that only studies with larger effects are published.”

>> Thanks for your correction. We added the results of the trim-and-fill procedure on body weight in the Risk of bias across studies of RESULTS. The trim-and-fill procedure was performed according to Duval and Tweedie (2000). We added this paper to the REFERENCES (No. 18). According to your advice, we also mention the meaning of publication bias in INTRODUCTION.

“Could you elaborate on the sentence at the end of this third paragraph on page 10 concerning the meta-regression finding DPA having significant time-dependent effects. Is this saying DPA levels change over time (presumably higher levels <= 24 weeks compared to >24 weeks) irrespective of FPG and body weight. If so is this suggesting the reason why the 'efficacy is time-dependent' on page 9 (first full paragraph) namely that since DPA levels decrease over time its effect decreases thus reducing differences in HbA1c, FPG and body weight between the DPA and control groups. Does this suggest a timeframe suggesting how long one should wait before topping up levels of DPA with further injections to control glucose levels?”

>> We are sorry for the incorrect use of ‘time-dependent’. The meta-regression was actually conducted on the changes of HbA1c, FPG, and body weight of the participants taking DAPA for different follow-up durations (12th, 24th, 48th, 50th, 102nd, 104th weeks). The analysis was not on the changes of DAPA level and when to top up the DAPA. Thanks for your question, the term ‘time-dependent effect’ has been amended to ‘long-term effect’.

“Page 26. Is it possible to add in the 95% confidence limits to the funnel plots in Figure 4 to give a stronger indication of publication bias? I would also add in the adjusted funnel plot with extra studies denoted as added by the trim and fill procedure described in the third paragraph on page 10.”

>> Thanks for your advice, according to which we have updated Figure 6. Both the 95% confidence intervals and adjusted funnel plots have been added.

“Pages 23, 24, 25, 27 and 28. STANDARDISED means difference would be more informative than ‘mean difference’ to describe the effects stated in the tables.”

>> Thanks for your advice. SMD is the difference in the means between the treatment and control divided by their SD. If the variability of the number of participants were artificially or accidentally reduced, according to Takeshimam et al. (2014), SMD would be overestimated. Hence, we used mean differences to describe the effects.

“In passing I would draw attention to Peters et al. (2010) who criticise the use of Egger’s test which is used here as a measure of publication bias.”

>> Thanks for your advice. Both funnel plots and the Egger’s test showed consistent results in our case. And we just used the results from the Egger’s test as an additional piece of evidence.

References:

Schmidt FL, Oh IS, Hayes TL. Fixed-versus random-effects models in meta-analysis: model properties and an empirical comparison of differences in results. British Journal of Mathematical and Statistical Psychology, 2009, 62(1): 97-128.

Takeshima N, Sozu T, Tajika A, et al. Which is more generalizable, powerful and interpretable in meta-analyses, mean difference or standardized mean difference? BMC Medical Research Methodology, 2014, 14(1): 30.

REVIEWER 2:

“Please state any competing interests or state ‘None declared’: None declared.”

>> Thanks for your advice. The statement of competing interests has been added.

“Major Comments
In the strengths and limitations the authors state “This study is the first meta-analysis to focus on the weight gain issue of dapagliflozin”. What does this meta-analysis add to those published previously such as Goring et al. Diabetes Obes Metab. 2013 Nov 14. doi: 10.1111/dom.12239?”

>> Thanks for your question. This research article was finished in last November, until when Goring et al. had not been published. In terms of differences in study design, Goring et al. took weight loss as an additional benefit and estimated the relative effect of dapagliflozn versus other treatment (not placebo) when added to only metformin. In our study, the objective was to explore the effect of dapagliflozin versus placebo, in combination with hypoglycemic drugs (not only metformin). We have

changed the statement as “This study is the first meta-analysis to focus on both the efficacy and body weight gain issue of dapagliflozin versus placebo in synergy with anti-diabetic drugs (not only metformin)” in the Strengths and limitations of this study (ARTICLE SUMMARY).

“The meta-analysis included a mixture of RCTs some of them were the same study reported at different time points, why the authors didn’t just included the final analysis with the longest follow up as these likely to be more meaningful clinically.”

>> Thank you for your question. We did perform the analysis with the longest follow-ups but reported in Sensitivity analysis (RESULTS). We are sorry for not explaining this well in manuscript; hence, we have added an explanation (“with final results”) in the third line, second paragraph of the Sensitivity analysis of RESULTS.

“Minor comments
In key messages, the authors state “Dapagliflozin as an add-on drug improves the control of HbA1c and FPG levels in type 2 diabetes mellitus (T2DM) patients with little risk of weight gain”; overall I agree but I do not think the results support the statement regarding little weight gain. Teh results suggest that Dapa treatment was associated with weight loss. In order to justify the authors statement they need to show us the risk of weight gain in Dapa treated patients which is not included in the current manuscript.”

>> Thanks for your correction. Individual RCTs indicated that some anti-diabetic drugs could increase body weight of T2DM patients. Our results demonstrated that anti-diabetic drugs in combination with dapagliflozin didn’t result in body weight gain. We have deleted our misleading term “risk” and amended the statements to ‘without body weight gain’.

“The issue of publication bias in regards to weight loss but not in regards to HbA1c in interesting. I agree that the publication bias is evident based on the funnel plot. Although this is significant statistically I am not sure whether a true bias exists. All the Dapa RCTs reported the change in glycaemic measures (HbA1c) and weight; in addition the primary outcome is usually change in HbA1c. I am not aware that any studies were conducted with weight being the primary outcome; hence it is unlikely for studies not to be published because of lack of effect on weight.”

>> Thanks for your helpful suggestion in improving the manuscript. We have added your suggestion to the fourth paragraph in DISSCUSION.