BMJ - Decision on
Manuscript ID
BMJ.2017.039000

Body:

Dear Prof. Tsapas

Manuscript ID BMJ.2017.039000 entitled "Closed-loop insulin therapy for outpatients with type 1 diabetes: a systematic review and meta-analysis"

12-Sep-2017
Dear Prof. Tsapas
BMJ.2017.039000 entitled "Closed-loop insulin therapy for outpatients with type 1 diabetes: a systematic review and meta-analysis"

Thank you for sending us your paper. We sent it for external peer review and discussed it at our Endgames meeting. We recognise its potential importance and relevance to general medical readers, but I am afraid that we have not yet been able to reach a final decision on it because several important aspects still need clarifying.

We hope that you will be willing and able to revise your paper according to the editors' comments and the peer review comments, which are included below, so that we will be in a better position to understand your article and decide whether The BMJ is the right journal for it. We are looking forward to reading the revised version and, we hope, reaching a decision.

Please note that resubmitting your manuscript does not guarantee eventual acceptance, and that your resubmission may be sent for further external peer review.

Please don't hesitate to contact me if you wish to discuss this further.

Yours sincerely,

Rubin Minhas
Dr Rubin Minhas
BMJ Associate Editor
rm1000@live.com

In your response please provide point by point replies to all the comments made by The BMJ editors and the external peer reviewers, explaining how you have dealt with them in the article.

EDITORS' COMMENTS TO AUTHOR:

REVIEWERS' COMMENTS TO AUTHOR:

Reviewer: 1

Comments for editor and author
In this systematic review and meta-analysis, Bekiari et al evaluated, as only outcome, % of time that sensor glucose level was within the near normoglycaemic range. They only considered randomized studies; they performed sensitivity analysis in order to reduce bias of the studies considered. The coverage of literature was accurate and up-to-date. The overall take-home message is clear and well substantiated.

However, there are two minor points:
1. it is not true that this is the first meta-analysis, as one meta-analysis on the same issue was published in 2011 (J Diabetes Sci Technol. 2011 Nov 1;5(6):1352-62). The authors should comment on what is new in their meta-analysis as opposed to the preceding meta-analysis.
2. The authors should look for consistency in their Forest Plots, as "favours closed" appera in some instances on the left and in other instances on the right

Confidential Comments to the Editor
(There are no comments.)

Reviewer: 2

Comments for editor and author
Closed-loop insulin therapy for outpatients with type 1 diabetes: a systematic review and meta-analysis – reviewed by Jennifer Hirst

This analysis has combined data from trials comparing closed-loop insulin therapy with other types of insulin therapy in populations with type 1 diabetes. The authors have combined data on percentage of time spend in the therapeutic range, time spent out of range and mean blood glucose as well as other outcomes. In all cases the closed-loop system outperformed conventional insulin therapy. This work was well conducted and methodologically sound.

Please clarify whether any adverse event data reported (in particular hypoglycaemic episodes) in the trials and how the numbers compared between the randomised groups. If there are sufficient data then a meta-analysis of this data should be included.

Please include a more comprehensive explanation of low blood glucose index and interpretation of the implications of a lower LBGI.
The numbers in Figure 1 are not quite consistent, please check.
There are 34 included studies reported in the text and Figure 1, but there are 37 studies listed in Table 1.
Some of the numbers in the text on page 5 do not add up to 34, for example 29 trials of single hormone and 8 trials of dual hormone exceeds the 34 included trials.
Risk of bias: "Most studies were deemed at high risk for bias due to incomplete outcome data.." This is misleading and should be revised; should it read something like "Of those studies at high risk for bias due to incomplete outcome data, …." ?
Figure 2 subgroup labelling is "single hormone CL" and" dual hormone CL", whereas the text on page 6 reports findings from "Closed-loop overnight" and "throughout24 hours".
The protocol states that data for area under the curve of glucose<3.5 mmol/l will be reported but it has not been reported in the manuscript. Please explain the reasons for this.
The table of included studies would benefit from a column showing length of follow-up.
Please report the mean time study participants were within therapeutic range at baseline if this data is available.
Discussion – comparison of the findings with the existing literature has not been included. Please add a section comparing this work with other research, in particular the previous systematic review published in 2014.

Confidential Comments to the Editor
(There are no comments.)

Reviewer: 3

Comments for editor and author

The outcome is % of time that sensor glucose level was within the near normoglycaemic range. Secondary outcomes are also defined on sensor glucose levels. I can't see (Figure 1) any trials being excluded because the outcome was not available in the control arm. I lack expertise on different types of insulin therapy (well, I'm reviewing for a generalist journal, not a diabetes journal). I can guess that when the control arm is "Sensor augmented pump therapy" then the outcomes, that depend on sensors, are available in the control arm. What about when the control arm is "Insulin pump therapy" or "Low Glucose Suspend"? Do these therapy protocols also use continuous glucose sensors? Or was continuous monitoring added by the trialists, for the sake of comparability? If the latter, did you pre-specify that you would exclude trials that did not use sensors in the control arm - or was it just that such trials did not arise during review?

Related to this: did the trials ensure equivalence of outcome measurement in the intervention and control patients? Specifically, was the frequency of sensor monitoring the same in both the treatment and control arms, to prevent confounding between the outcomes and the intervention? I can't see any mention or exploration of this in the manuscript - apologies if I have overlooked it.

Given that this has been submitted to a generalist journal, rather than a diabetes journal, explanation of the different types of therapy would be welcome.

Please could you supply a reference for "counter-enhanced" funnel plots? Standard funnel plots are centered on the main effect; this one appears to be centred on zero. Visual inspection of the funnel plot is mentioned in the manuscript; are you aware that there is empirical evidence that visual inspection is too subjective to be useful? See Terrin, Schmid & Lau, 2005, J Clin Epi 58:894-901. Based on this it is more useful to include the p-value than "visual inspection". The results refer to this p-value as "significant publication bias", but it is more strictly correct to describe this as "evidence of small-study effects" than to assume that small-study effects are always due to publication bias.

The statistical heterogeneity is high, but that is not uncommon in evaluations of complex interventions. The authors have explored some candidate explanations for heterogeneity: single vs dual hormone therapy; overnight vs 24hr use of the system. Both of these are characteristics of the intervention. I don't wish to encourage the authors to explore too many hypotheses, on this number of studies, but is it worth also exploring whether characteristics of the comparator (SAP vs insulin pump) explain some of the heterogeneity between studies?

Confidential Comments to the Editor
(There are no comments.)

Reviewer: 4

Comments for editor and author
Review of “Closed-loop insulin therapy for outpatients with type 1 diabetes: a systematic review and meta-analysis” by Bekiari et al.

Thank you for the chance to review this comprehensive and important paper.

1) The topic is absolutely relevant for BMJ's readers. Everyone working with type 1 diabetes is looking for the size of overall treatment improvements, with closed loop therapy systems. All because we are looking for anything, which can diminish the development of diabetes late complications in our patients. Secondly, closed-loop insulin therapy most probably will impact positively the patient burden in decision making on insulin dosing; however, this was not the topic for the current review.
2) This systematic review and metaanalysis is done according to guidelines and best principles for such kind of research – and I cannot find anything which should have been done methodological differently.
3) Regarding the primary and secondary outcomes; comparisons of time in target (and above/below target), I agree that the differences between closed loop and control treatment are the main outcomes and should be reported in text and figures. However, data on the actual time spent in target during closed loop and control treatment in the different studies (and above/ below targets) would be of additional values for the reader, given in result text only. I suggest roughly statements on how these percentages vary between studies for closed loop and control treatments.
4) I suggest adding the actual values for sensor glucose and HbA1c in text as well (and not only changes).
5) For the discussion I appreciate its precise and short form pointing on better reporting and broader patient selection in further closed loop studies.
6) In Table 2, I will suggest for the readers less familiar with parameters as time in target etc. to emphasize in legends that here is given differences between the two treatments compared (and not actual values). Also in all figure legends, I suggest to add, mean difference between closed loop treatment and control treatment.
7) If space is a problem for the journal, I will suggest Fig 8 to be moved to appendix rather than in the main paper and results given in main text only.

Confidential Comments to the Editor
(There are no comments.)

Reviewer: 5

Comments for editor and author
Thank you for the opportunity to review this interesting paper. The authors have clearly done a lot of work to identify, appraise and snythesise the existing evidence in this field. I have reviewed this from a statistical perspective, and have some comments for the authors to address going forward:
1) I worry about the quality of the individual studies. Many of them have just 10-30 patients. I doubt randomisation could have balanced the groups in such a short sample. Did the individual studies have balance at baseline? Was adjustment for any imbalance undertaken? When taking results from published studies, it is hard to overcome issues of baseline imbalance without the IPD. Even then, adjustment for non-recorded baseline variables is not possible. Can the authors reassure the BMJ that there syntheses are meaningful? Was this issue accounted for in the risk of bias classification?

2) I am also concerned about the primary outcome definition: % of time in a normal range. Is this meaningful clinically? Is 'normal' well defined. For example, for a value just outside the range (e.g. 10.01mmol/l), why should this be abnormal when a neighbouring (almost identical) value (e.g. 9.99mmol/l) that happens to fall inside the range is classed as normal. Of course, this may be how the primary studies recorded the outcome, but this classification has consequences for interpretation of

the meta-analysis results. I find the outcome uncomfortable. Similallry, others like % of time > 10mmol/l, or below 3.9mmol/l.

3) I am pleased that the authors use a random effects analysis. However, they do not say what estimation method was used.

4) In relation to this, there is increasing evidence that the uncertainty in heterogeneity estimates should be accounted for in the derivation of CIs. See refs below [1, 2]. For example, the Hartung Knapp correction works generally quite well.

5) With large heterogeneity as observed here, 95% prediction intervals can be helpful to summarise the range of effects across settings better than the summary effect itself. i.e. the average effect is perhaps not so meaningful. The authors might consider an approximate prediction interval to address this. [3]

6) Publication bias assessments should be better explained as assessments of small study effects [4]

7) In the abstract, please state how many trials were at low risk of bias

8) Referring back to the outcome, I see a range of secondary outcomes. But was the trend in glucose levels not summarised? E.g. in some studies, were the repeated measures of glucose not modelled, and could these have not been synthesised (e.g. mean trend)?

9) STATA should be Stata

10) An I-squared > 50% does not necessarily indicate high heterogeneity. Please see Rucker [5]. It depends on the size of the studies. An actual estimate of tau-squared (between-study variance) is preferable.

11) The authors say they explored risk of bias using a funnel plot and Egger's test. But why does this relate to risk of bias? All studies could have high risk of bias but the funnel plot may be symmetrical. Perhaps the 'risk of bias' is misleading language. Do you mean risk of bias of the original studies or risk of bias in the summary effect due to small study effects? I think the latter perhaps? Anyway, there is some confusion with the risk of bias tool.

12) Most studies were high risk of bias because "they reported median instead of mean values or reported results that required extensive use of imputation methods to be used in meta-analyses". Why is median is worse than mean. Indeed, if the % time in normal level is skewed, the median may be preferred. Also, if the imputation methods are not reliable, then why does it say 'appropriate formulas to calculate mean and variance" in the methods. This warrants further explanation please.

13) Contour enhanced funnel plot is not mentioned in the methods

14) In the results, when giving a summary result I suggest saying 'summary' or 'average' effect explicitly. Also, in the brackets please give the number of studies next to each m-a result. Some description of the amount of heterogeneity is also warranted in the primary and secondary outcome results sections. This is where a prediction interval may be warranted, to summarise the range of effects across settings.

15) Please include the results for low risk of bias in the main paper, not just appendix.

16) I like the discussion about the limitations of the existing studies, and what needs to improve in terms of reporting, included populations, sample sizes, and follow-up length. However, the abstract never mentions these limitations or recommendations, and this should be addressed to give a better-rounded summary of the evidence found. I would also include the outcome definition as a major limitation and something for new trials to address.

17) A similar comment applies for the 'what this study adds'.

18) Table 2 – are these summary meta-analysis results? If so, we have no idea of the number of studies, amount of heterogeneity, etc Also, if the aim is to compare single and dual therapy, then a meta-regression should have been conducted and the difference in the groups formally estimated for each outcome.

I hope these comments are helpful the authors to revise their article and improve their work further.
Best wishes, Richard Riley
Reference List

1. Cornell JE, Mulrow CD, Localio R, Stack CB, Meibohm AR, Guallar E, Goodman SN. Random-effects meta-analysis of inconsistent effects: a time for change. Ann Intern Med 2014; 160: 267-270.
2. Partlett C, Riley RD. Random effects meta-analysis: Coverage performance of 95% confidence and prediction intervals following REML estimation. Stat Med 2016.
3. Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. BMJ 2011; 342: d549.
4. Sterne JAC, Sutton AJ, Ioannidis JPA, Terrin N, Jones DR, Lau J, Carpenter J, Rücker G, Harbord RM, Schmid CH, Tetzlaff JT, Deeks JJ, Peters J, Macaskill P, Schwarzer G, Duval S, Altman DG, Moher D, Higgins JPT. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ 2011; 342: d4002.
5. Rucker G, Schwarzer G, Carpenter JR, Schumacher M. Undue reliance on I(2) in assessing heterogeneity may mislead. BMC Med Res Methodol 2008; 8: 79.