PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Comparative effectiveness of sodium-glucose co-transporter 2 inhibitors for controlling hyperglycemia in patients with type 2 diabetes: protocol for a systematic review and network meta-analysis
AUTHORS	Chen, Min; Xie, Chun-Guang; Gao, Hong; Zheng, Hui; Chen, Qin; Fang, Jianqiao

VERSION 1 - REVIEW

REVIEWER	William L. Baker University of Connecticut School of Pharmacy Storrs, Connecticut, USA
REVIEW RETURNED	23-Oct-2015

GENERAL COMMENTS	While I question the ultimate utility of the information provided in this paper beyond what is available on PROSPERO, it is adequately written (the journals editors can fix the English grammar errors). I have no edits to recommend.

REVIEWER	Norman Waugh Warwick Medical School, UK I am corresponding author of a network meta-analysis of the same six SGLT2 inhibitors submitted to BMJ Open on 17th July. We await a decision. We have not yet received any referees reports. However our NMA has a more restricted scope (just monotherapy and dual therapy, and not safety). Comments to editor. I have said in a tick box below that I would review a revised version but the revisions necessary and minor and a further review will probably not be necessary. Incidentally, we are now seeing a lot of reviews coming from China. Initially some were poor but quality has improved
REVIEW RETURNED	27-Oct-2015

GENERAL COMMENTS	Ref report BMJ Open 2015-010252 For authors This is a good protocol and I have few comments, mostly minor. Abstract, Introduction. A network meta-analysis of all six SGLT2 inhibitors was submitted to BMJ Open on 17th July 2015 but covers only relative effectiveness in monotherapy and dual therapy, and does not cover safety. No decision from BMJ Open has been received by 27th October. Methods, date of searches is reported as August in Abstract but November in Methods page 7. Since the protocol was written, the Empagliflozin Outcomes study has been published, reporting a reduction in cardiovascular events

	Zinman et al NEJM September 2015. Introduction Line 2, I suggest you specify the level of HbA1c used for diagnosis. Page 5, 3 lines from foot of page. Some studies have shown larger reductions in HbA1c with larger doses of canagliflozin so this statement applies only to the Rosenstock trial. Methods Page 6, last sentence raises an issue. Does publication of the protocol in PROSPERO reduce the value of a subsequent journal version of the protocol? Does it represent dual publication? The journal version has more details of background. Page 7, para 1, data sources, mentions searches for abstracts. How will these be used? They won't provide enough data for quality assessment. Will trials reported only in conference abstracts be included in meta-analyses? You might handle such abstracts differently if trial protocols have been published. Otherwise caution is needed. Page 7, study design. RCTs only? What about longer-term observational studies for adverse events. Such as recent reports on DKA. Data sources. You could add the reports used to support the NICE appraisals of dapagliflozin, canagliflozin and empagliflozin. The three Evidence Review Group reports for use in combination therapy are on the NICE (www.nice.org.uk) website and the assessment group report on all three in monotherapy will be in two months time. These reports sometimes have material not in the published papers and can also be used for checking completeness of capture of trials. Page 8, first sentence. The duration periods don't seem optimal given the insulin-independent mechanism of action. I suggest having longer periods including > 10 years. Page 8, para 2. I suggest having a minimum duration of treatment of at least 24 months. Page 10, para 2 should say "Cochrane Collaboration" not library. Page 10, and page 11, data analysis. You need to compare effect sizes separately in four situations: monotherapy, dual therapy, triple or quadruple therapy, and combinations with insulin. Not just monotherapy and add-ons as said on page 11. Page 11, para 1, line 7. What level of heterogeneity will mean that meta-analysis is inappropriate? Pages 12-13, sensitivity analyses, and risk of bias scores. Will you report which trials are sponsored by the manufacturers of the drugs? I expect they all will be but you can look for independent RCTs. Discussion Trials compared canagliflozin 100 mg with canagliflozin 300mg, and empagliflozin 10mg with empagliflozin 25mg in patients who had not been on SGLT2 inhibitors before. But the licences say that the larger doses should only be used in people who have tolerated the smaller doses but have not had an adequate effect in HbA1c. People who don't respond adequately to the stating dose may be poor responders to SGLT2 inhibitors, so the effects of canagliflozin 300 and empagliflozin 25 may be less than seen in the trials. How will you deal with that? Will you include trials of fixed dose combinations of an SGLT2 inhibitor with a DPP4 inhibitor such as linagliptin? Such combinations may help with compliance with medication.

REVIEWER	Lise Lotte Gluud Copenhagen University Hospital Hvidovre, Denmark
REVIEW RETURNED	10-Nov-2015

GENERAL COMMENTS	This paper describes a protocol for a systematic review evaluating sodium-glucose co-transporter 2 inhibitors (SGLT2) for people with type 2 diabetes. The authors plan to conduct network meta-analyses. Although, the objective and design seems straightforward, I do have a few comments. The design and analyses in meta-analysis networks should be very carefully planned. At present, there are no RCTs comparing the different SGLT2, so all comparisons will be indirect. The different trials include a large number of co-interventions, which will affect the estimated intervention effect. The dose of the intervention is essential and should always Also, the patient inclusion criteria are important since we know that the relative effect of the intervention will depend on aspects such as inclusion of people with renal disease. Other important aspects such as the diabetes duration and the level of HbA1c at baseline. The ethnicity of included participants is also important and certain drugs (such as luseogliflozin) are tested in trials from Asia. This will increase heterogeneity and make it very difficult to evaluate the overall conclusions. The authors put a lot of focus on the ranking of the different SGLT2. However, before making any conclusions regarding the relative effect, the analyses should carefully account for characteristics of participants. I disagree that it will be reasonable to include crossover trials in the analyses considering the nature of the primary outcome. If crossover trials are included, I would suggest to restrict the analyses to the first treatment period. The analyses and design of the review put much emphasis on HbA1c. However, we know that there are important clinical outcomes which should be included in the evaluation. One important assessment is the risk of adverse events. Recent evidence highlights that serious adverse events including ketoacidosis and cancer should be considered in the analysis. The information may not be available in RCTs. I would therefore suggest including serious adverse events as a primary outcome and to include observational studies in the assessment. Based on the wording, it appears that the inclusion of trials will be based on reported outcomes. This approach is debatable as it will increase the risk of reporting bias. In the assessment of outcomes, the authors need to include additional information. For example, we don't know if the authors will use HbA1c levels or %change. The authors have made a judgement that they only plan to include randomization and blinding in the assessment of the overall quality. However, this approach will exclude important bias domains such as attrition bias and reporting bias. I am not convinced that such an analysis will provide the clearest picture of the overall risk of bias. The choice to use only three domains in the overall analysis means that the strategy does not follow the recommendations used in the Cochrane Handbook. This should be clarified and if possible, the authors should include a reference that will support their method. I was very surprised to see that the authors plan to use a frequentist network rather than a Baysean network. I would suggest writing a paragraph in the background section about the choice. I found the starting data somewhat confusing. In PROSPERO (registered September 7), the authors write that they have not yet searched the literature for eligible trials, but in the submitted protocol, the authors write that they will search several databased

	from the date of inception to August 2015. Why was this particular date chosen as a cut-off? The reported outcomes in PROSPERO are Primary outcomes The primary outcome of this study will be the changes of haemoglobin A1c from baseline compared to baseline. Secondary outcomes The secondary outcome will include proportion of participants who achieve a normal level of blood glucose, changes of body weight, and changes of blood pressure and incidence of any adverse events (especially urinary and genital tract infections and cardiovascular events). In the submitted protocol, the authors write that 'The primary outcome of this study will be the mean change from baseline in HbA1c. The secondary outcome will include proportion of participants with normal HbA1c, mean change from baseline in body weight, mean change from baseline in blood pressure and incidence of adverse events (especially urinary and genital tract infections and cardiovascular events).' I would suggest that the authors revise the registered information in order to ensure that the information is the same. Do the authors plan to evaluate The proportion of patients who achieve a normal blood glucose or proportion with 'normal' HbA1c?

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
William L. Baker
University of Connecticut School of Pharmacy
Storrs, Connecticut, USA
Please leave your comments for the authors below
(1) While I question the ultimate utility of the information provided in this paper beyond what is available on PROSPERO, it is adequately written (the journals editors can fix the English grammar errors). I have no edits to recommend.
Response: I appreciate the comments from prof. Baker. Our protocol provides more information than what is available on PROSPERO, especially the information about data analysis. We have checked and revised the manuscript for grammar mistakes, please read the article.

Reviewer: 2
Norman Waugh
Warwick Medical School, UK

(1)This is a good protocol and I have few comments, mostly minor.
Abstract, Introduction. A network meta-analysis of all six SGLT2 inhibitors was submitted to BMJ Open on 17th July 2015 but covers only relative effectiveness in monotherapy and dual therapy, and does not cover safety. No decision from BMJ Open has been received by 27th October.
Response: Thank you for the information, I did not know that when we submitted our protocol. Our meta-analysis will cover the assessment of safety outcomes of the six SGLT-2 inhibitors. We change the background information in the abstract.
(2)Methods, date of searches is reported as August in Abstract but November in Methods page 7.
Response: The initial of the study was postponed to November 2015, we revise the abstract, see page 2.

(3)Since the protocol was written, the Empagliflozin Outcomes study has been published, reporting a

reduction in cardiovascular events Zinman et al NEJM September 2015.
Response: Thank you for noticing us. We found this article after submitting our protocol. We cite this article now and revise the inclusion criteria and the wording in strengths and limitation section. See page 3 and 7.

(4)Introduction
Line 2, I suggest you specify the level of HbA1c used for diagnosis.
Response: According to the guideline (released by the American diabetes association (ADA) and the European association for the study of diabetes (EASD)), we added this information in criteria for screening studies, in page 7.

(5)Page 5, 3 lines from foot of page. Some studies have shown larger reductions in HbA1c with larger doses of canagliflozin so this statement applies only to the Rosenstock trial.
Response: I agree with you on this point. We have corrected this statement as" Rosenstock and colleagues found that 50mg canagliflozin worked better in lowering HbA1c than 200mg canagliflozin." See page 4.
Methods
(6)Page 6, last sentence raises an issue. Does publication of the protocol in PROSPERO reduce the value of a subsequent journal version of the protocol? Does it represent dual publication? The journal version has more details of background.
Response: It is a very good point. Publication of the protocol in PROSPERO may partly reduce the value of a subsequent journal version of the protocol, since some of the content will be the same in the registered protocol and the journal version. However, it is necessary to publish the protocol in a journal version. First, publishing study protocols enable researchers to stay up to date in their study fields, which is one of the aims in publishing protocols in BMJ Open. If we had not tried to publish this study protocol in BMJ Open, I would not know that a network meta-analysis of all the six SGLT2 inhibitors was submitted to BMJ Open on 17th July 2015, and this information helps us improving our protocol. Second, the protocol in a journal version will provide more details than the registration protocol. It will contain information like how we consider in designing this protocol, which was not introduced in the PROSPERO. And this is not a dual publication, as you could see that changes are made in the journal version of the protocol, it is different from the one in the PROSPERO.
(7)Page 7, para 1, data sources, mentions searches for abstracts. How will these be used? They won't provide enough data for quality assessment. Will trials reported only in conference abstracts be included in meta-analyses? You might handle such abstracts differently if trial protocols have been published. Otherwise caution is needed.
Response: I agree with you. Trials reported in conference abstracts will not be included in meta-analysis, they will be used in a narrative review. We add an explanation in the data source section, see page 6.
(8)Page 7, study design. RCTs only? What about longer-term observational studies for adverse events. Such as recent reports on DKA.
Response: This is a very good point. We will include long-term observation studies for adverse events. We revised the selection criteria in page 7.
(9)Data sources. You could add the reports used to support the NICE appraisals of dapagliflozin, canagliflozin and empagliflozin. The three Evidence Review Group reports for use in combination therapy are on the NICE (www.nice.org.uk) website and the assessment group report on all three in monotherapy will be in two months time. These reports sometimes have material not in the published papers and can also be used for checking completeness of capture of trials.
Response: It is a very good suggestion, we add this information in page 6.
(10) Page 8, first sentence. The duration periods don't seem optimal given the insulin-independent mechanism of action. I suggest having longer periods including > 10 years.
Response: We revised the division of the duration periods in page 8 according to your suggestion.
(11) Page 8, para 2. I suggest having a minimum duration of treatment of at least 24 months.

Response: We revised according to the results of your review and the others, we will include trials with a minimum duration of 12 weeks, see page 8.
(12)Page 10, para 2 should say “Cochrane Collaboration” not library.
Response: See page 10.
(13)Page 10, and page 11, data analysis. You need to compare effect sizes separately in four situations: monotherapy, dual therapy, triple or quadruple therapy, and combinations with insulin. Not just monotherapy and add-ons as said on page 11.
Response: It is a very good point, we revised in page 10 and 11.
(14)Page 11, para 1, line 7. What level of heterogeneity will mean that meta-analysis is inappropriate?
Response: An I2>50% will be considered as meaningful heterogeneity, however, we will not make a choice of fixed or random effect model based on I2 statistic alone, we will run meta-regressions and find out potential cofounders and exclude these cofounders to see it solve the heterogeneity problem. If the problem is not solved after subgroup analysis and meta-regression, we will not perform meta-analysis. We revise this in page 10 and 11.
(15)Pages 12-13, sensitivity analyses, and risk of bias scores. Will you report which trials are sponsored by the manufacturers of the drugs? I expect they all will be but you can look for independent RCTs.
Response: It is a very good idea, we will report which trials are sponsored by the manufacturers of the drugs, see page 13, the sensitivity analysis section.

(16)Discussion
Trials compared canagliflozin 100 mg with canagliflozin 300mg, and empagliflozin 10mg with empagliflozin 25mg in patients who had not been on SGLT2 inhibitors before. But the licences say that the larger doses should only be used in people who have tolerated the smaller doses but have not had an adequate effect in HbA1c. People who don’t respond adequately to the stating dose may be poor responders to SGLT2 inhibitors, so the effects of canagliflozin 300 and empagliflozin 25 may be less than seen in the trials. How will you deal with that?
Will you include trials of fixed dose combinations of an SGLT2 inhibitor with a DPP4 inhibitor such as linagliptin? Such combinations may help with compliance with medication.
Response: That is a very good question, and you have provided a good answer for it. Patients using the larger doses may be poor responders to SGLT2 inhibitors. We will include trials of fixed dose combinations of an SGLT2 inhibitor with a DPP4 inhibitor (patients in these trials may also be poor responders, so they used dual therapy). We will compared all monotherapy, dual therapy, triple or quadruple therapy in a network meta-analysis, through this method, we could observe which therapy achieve the best effect in poor responders, using larger doses or using combinations such as SGLT2 + DPP4. See page 12 and 14.

Reviewer: 3
Lise Lotte Gluud
Copenhagen University Hospital Hvidovre, Denmark
Please leave your comments for the authors below
(1) This paper describes a protocol for a systematic review evaluating sodium-glucose co-transporter 2 inhibitors (SGLT2) for people with type 2 diabetes. The authors plan to conduct network meta-analyses. Although, the objective and design seems straightforward, I do have a few comments. The design and analyses in meta-analysis networks should be very carefully planned.
At present, there are no RCTs comparing the different SGLT2, so all comparisons will be indirect. The different trials include a large number of co-interventions, which will affect the estimated intervention effect. The dose of the intervention is essential and should always…
Response: I agree that the different trials use a large number of co-intervention and that ranging doses of SGLT2 inhibitors are used and compared in the different trials. Therefore, we compare effect sizes separately in four situations: monotherapy, dual therapy, triple or quadruple therapy. And the 6

SGLT2 inhibitors will be compared in pairs, with each of the SGLT2 inhibitors being administered in different doses (eg, 10mg empagliflozin vs. 25mg empagliflozin vs. 100mg dapagliflozin). We provide doses ranges of the 6 SGLT2 inhibitors and the use of co-interventions in the interventions section. See page 8.

(2) Also, the patient inclusion criteria are important since we know that the relative effect of the intervention will depend on aspects such as inclusion of people with renal disease. Other important aspects such as the diabetes duration and the level of HbA1c at baseline. The ethnicity of included participants is also important and certain drugs (such as luseogliflozin) are tested in trials from Asia. This will increase heterogeneity and make it very difficult to evaluate the overall conclusions.
Response: I agree. We revise the inclusion criteria according to your suggestions. Please see page 7-8.

(3) The authors put a lot of focus on the ranking of the different SGLT2. However, before making any conclusions regarding the relative effect, the analyses should carefully account for characteristics of participants.
Response: It is a good point. We revise the wording and will run subgroup analyses considering the duration of diabetes and ethnicity in the included participants. Please see the statistical section, page 11.
(4) I disagree that it will be reasonable to include crossover trials in the analyses considering the nature of the primary outcome. If crossover trials are included, I would suggest to restrict the analyses to the first treatment period.
Response: We exclude crossover trials as you suggest, see page 7.
(5) The analyses and design of the review put much emphasis on HbA1c. However, we know that there are important clinical outcomes which should be included in the evaluation. One important assessment is the risk of adverse events. Recent evidence highlights that serious adverse events including ketoacidosis and cancer should be considered in the analysis. The information may not be available in RCTs. I would therefore suggest including serious adverse events as a primary outcome and to include observational studies in the assessment.
Response: I agree. We revise the design to include observation studies that assess the adverse events (including events of cardiovascular disease, ketoacidosis and cancer) for using SGLT2. And we define the outcome on HbA1c more explicit. Please see page 8.
(6) Based on the wording, it appears that the inclusion of trials will be based on reported outcomes. This approach is debatable as it will increase the risk of reporting bias.
Response: No, we will not exclude trials based on reported outcomes, it will greatly increase the risk of reporting bias. We cause misunderstanding in writing the inclusion criteria section, we revise this section, see page 8.

(7) In the assessment of outcomes, the authors need to include additional information. For example, we don't know if the authors will use HbA1c levels or %change.
Response: We revise the outcome measurements section, see page 8.
(8)The authors have made a judgement that they only plan to include randomization and blinding in the assessment of the overall quality. However, this approach will exclude important bias domains such as attrition bias and reporting bias. I am not convinced that such an analysis will provide the clearest picture of the overall risk of bias. The choice to use only three domains in the overall analysis means that the strategy does not follow the recommendations used in the Cochrane Handbook. This should be clarified and if possible, the authors should include a reference that will support their method.
Response: I agree with on this point. I revise the section of quality assessment and will assess the overall quality with inclusion of all the 6 domains, see page 10.
(9) I was very surprised to see that the authors plan to use a frequentist network rather than a Bayesian network. I would suggest writing a paragraph in the background section about the choice.

Response: We add a section to introduce the method we used, see page 5.
(10) I found the starting data somewhat confusing. In PROSPERO (registered September 7), the authors write that they have not yet searched the literature for eligible trials, but in the submitted protocol, the authors write that they will search several databased from the date of inception to August 2015. Why was this particular date chosen as a cut-off?
Response: We planed to start in August, but actually started in November because of unavailability of some of the authors during Aug to Oct. We make changes to PROSPERO to keep it up to date. Please see the abstract and data source section in page 2 and 6, respectively.
(11)The reported outcomes in PROSPERO are
Primary outcomes
The primary outcome of this study will be the changes of haemoglobin A1c from baseline compared to baseline.
Secondary outcomes
The secondary outcome will include proportion of participants who achieve a normal level of blood glucose, changes of body weight, and changes of blood pressure and incidence of any adverse events (especially urinary and genital tract infections and cardiovascular events).
In the submitted protocol, the authors write that 'The primary outcome of this study will be the mean change from baseline in HbA1c. The secondary outcome will include proportion of participants with normal HbA1c, mean change from baseline in body weight, mean change from baseline in blood pressure and incidence of adverse events (especially urinary and genital tract infections and cardiovascular events).'
I would suggest that the authors revise the registered information in order to ensure that the information is the same. Do the authors plan to evaluate The proportion of patients who achieve a normal blood glucose or proportion with 'normal' HbA1c?
Response: We made changes to the protocol during further discussion and writing process of the submitted protocol to BMJ Open. We update the information in PROSPERO now. By saying patients with normal HbA1c, we mean a level of HbA1c <7%. We revise the protocol to make it more explicit.

VERSION 2 – REVIEW

REVIEWER	Norman Waugh University of Warwick, UK
REVIEW RETURNED	16-Dec-2015

GENERAL COMMENTS	Thanks for making all your responses, which are fine. I have three last comments. Firstly, you might consider doing subgroup or sensitivity analyses excluding some studies carried out in East Asian populations where baseline BMI is much lower than in most other populations, or where the baseline HbA1c was low, as in the Kaku trial of dapagliflozin. Secondly, you should also note differences in the placebo groups, where HbA1c fell in some dapagliflozin trials but rose in some canagliflozin ones. Lastly, be quick - there are lots of reviews of the flozins being published.

REVIEWER	Lise Gluud Gastrounit, Hvidovre Hospital, Copenhagen University, Denmark
REVIEW RETURNED	28-Dec-2015

GENERAL COMMENTS	The reviewer completed the checklist but made no further comments.