PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Comparative safety of the sodium glucose co-transporter 2 (SGLT2) inhibitors: A systematic review and meta-analysis.
AUTHORS	Donnan, Jennifer; Grandy, Catherine; Chibrikov, Eugene; Marra, Carlo; Aubrey-Bassler, Kris; Johnston, Karissa; Swab, Michelle; Hache, Jenna; Curnew, Daniel; Hai, van Nguyen; Gamble, John Michael

VERSION 1 – REVIEW

REVIEWER	George Bakris University of Chicago Medicine
REVIEW RETURNED	16-Mar-2018

GENERAL COMMENTS	The author attempts to address a “knowledge gap” surrounding post-market serious safety outcomes of (SGLT2) inhibitors identified by the FDA and EMA. To do this they conducted a systematic review and meta-analysis of randomized controlled trials (RCT) and used random effects models to estimate pooled relative risks. The primary outcomes were: Acute kidney injury (AKI), diabetic ketoacidosis (DKA), urinary tract infections (UTI), bone fractures, and amputations. They found 99/1865 citations meeting their criteria. When compared to placebo, SGLT2 inhibitors were found to be significantly protective against AKI, while no difference was found for DKA, UTI, or bone fracture. No increased risk for either outcome was found when compared to active controls, and no studies reported on amputations. Sub-group analysis did show an increased risk of UTI with dapagliflozin only, but no other analysis supported an increased risk of AKI, DKA, UTI, or fracture. The authors conclude that current evidence does not suggest an increased risk of harm with SGLT2 inhibitors as a class over placebo or active comparators with respect to the AKI, DKA, UTI or fracture. However, wide confidence intervals for many comparisons suggest limited precision, and therefore clinically important adverse events cannot be ruled out. This is a thoughtful and well-done study. It has several strengths and some weaknesses. On the positive side the study is comprehensive on a timely subject. However, they covered events reported as having occurred, there is no information regarding the pathophysiology surrounding the events, i.e. volume depletion as a cause of AKI and urinary incontinence with UTIs. Moreover, certain outcomes are inadequately characterized within study reports. So, the major limitation of this analysis is its dependence the description of other trials. Beyond these comments this reviewer would caution about comments based on limited data, i.e. Dapa being accused of more

	UTIs than the others-that is based strictly on observation of some studies and is not backed by any pathophysiological reason for this to occur. Hence, it is a “nonstarter”. The authors should consider adding under each of the uncommon untoward effects a short paragraph alluding a possible pathophysiological basis for this certainly exist for all the events noted. That will add credibility to the paper.

REVIEWER	Marwan Saad Division of Cardiovascular Medicine, University of Arkansas for Medical Sciences, Little Rock, AR, United States
REVIEW RETURNED	10-Apr-2018

GENERAL COMMENTS	Donnan et al conducted a systematic review and meta-analysis aiming to evaluate the safety of SGLT2 inhibitors. They included 99 RCTs in their analysis and concluded that SGLT2 inhibitors are not associated with increased risk of harm compared with placebo or active comparators with respect to the AKI, DKA, UTI or fracture. The study is well conducted and written. However, some concerns to be addressed include: 1. The search strategy of the current study was until July 2017. It would be of importance to updated the search to evaluate if more recent studies have been published from July 2017 till now. 2. The authors mentioned that no data were available regarding the risk of amputations with SGLT-2 inhibitors. However, in a recent study published in New England Journal of Medicine in 2017 (N Engl J Med 2017; 377:644-657), data from CANVAS and CANVAS-R trials showed significant increased risk of amputations (up to 2-fold increase) with Canagliflozin compared to placebo. Furthermore, in a study by Monami et al in Acta Diabetol (2017) 54:411–413, 17 trials other than the CANVAS reported data regarding toe amputations. Authors should review the CANVAS, CANVAS-R, paper by Monami et al, as well as clinicaltrials.gov for these trials and include data comparing amputations with SGLT-2 inhibitors versus other comparators to the current paper. 3. Authors should report the weighted mean follow-up duration for each outcome reported, with the weight being the size of the population of each trial. 4. Authors reported data regarding UTI, however no data were reported regarding genital infections, which were reported as serious adverse effect of such class in many studies and meta-analyses (Saad et al, Int J Cardiol. 2017;228:352-358). Authors may report the adverse event of genital infections, or at least explain in their discussion the data in the literature about it. 5. In the introduction, authors mentioned “No meta-analysis on the risk of DKA currently exists.”, however a prior meta-analysis that evaluated the cardiovascular outcomes with SGLT-2 inhibitors evaluated the risk of DKA as well and found no significant increase compared with placebo (Saad et al, Int J Cardiol. 2017;228:352-358). This makes the above sentence incorrect. 6. Authors decided to include different SGLT2 inhibitors as an eligible comparator. This may cause a confounding effect in the reported outcomes. It would be more accurate to limit the

	comparator arm to placebo or other non- SGLT2 inhibitor antidiabetic medications. 7. For assessment of risk of bias, authors utilized the Cochrane Collaboration domain-based tool for assessing the risk of bias at the level of the included trial. It is recommended to further perform assessment of risk of bias at the level of the reported outcomes using GRADE tool as recommended by the Cochrane book for meta-analysis. 8. References need to be revised. For example in the discussion the following sentence has a wrong reference "However, a meta-analysis published in 2017, which is the largest to date, included 77 RCTs representing 50,820 patients and found no increased risk of UTIs in SGLT2 inhibitor users (RR 1.05; 95% CI 0.98-1.12).[16]". I believe the reference should be 14 rather than 16. Please revise the reference list for accuracy. 9. This sentence in the introduction needs language revision "They identified 73 cases of DKA and 19 cases of life-threatening infections that originated as a UTI, had been identified in patients taking a SGLT2 inhibitor."

REVIEWER	Jonathan R Treadwell ECRI Institute, USA
REVIEW RETURNED	07-May-2018

GENERAL COMMENTS	Statistical review of BMJ manuscript: "Comparative safety of the sodium glucose co-transporter 2 (SGLT2) inhibitors: A systematic review and meta-analysis" Jonathan R Treadwell Ph.D., ECRI Institute, May 2018 When speaking of non-significant differences, the authors should not lump them all together as "not suggesting a difference" or other vague wording. The problem is that some n.s. findings are clearly indicative that there is no important difference (e.g., your analysis of UTI) whereas other n.s. findings are simply inconclusive due to wide CIs (e.g., your analyses of DKA and bone fractures). While it is technically true that most your analyses "do not suggest an increased risk of harm", many readers will mistakenly interpret that as "evidence shows that there is no increased risk of harm". To avoid this reader misinterpretation, you should clearly delineate which n.s. outcomes are inconclusive, and which n.s. outcomes demonstrate that there is no effect (via a narrow CI). For example, consider the data on UTI for specific medications vs active comparators. When you analyzed empagliflozin, the CI was quite narrow (0.85 to 1.19) indicating that there is no difference. By contrast, when you analyzed other SGLT2's, the data are inconclusive (CI 0.31 to 2.15). These should not be presented identically as "no suggestion of a difference" or "no evidence of an effect". Clearly, with empagliflozin, there is enough evidence to assure us of no added UTI risk. The same cannot be said of other SGLT2's. Page 6 line 19. Change "primary" to "only". Page 7 line 26. You conducted multiple meta-analyses, not just one.

	Page 7 line 26. Please specify whether the random effects meta-analysis was the Dersimonian and Laird approach, and if so, cite that paper. Page 7 line 29. If a trial compared two SGLT2 medications, should it really be considered together with a trial that compared one SGLT2 medication to a non-SFLT2 medication? The former does not measure the added risk of using an SGLT2, but the latter does. Page 7 line 32. Tau is a more direct measure of heterogeneity than I2. See the article by Rucker in BMC Med Res Methodol 8(1) p79. The problem with I2 is that it depends on the Ns. Try it yourself: take a meta-analysis and look at I2. Now triple the Ns, but keep all the effect sizes the same. See how I2 increased? Yeah. That's not good. Note that tau would not have increased, since it more purely measures the extent to which effect sizes differed. Page 7 line 34. You say you used 75% as a threshold for "significant" heterogeneity. Do you have a reference for that? I know that 50% is often the threshold used, citing the original I2 paper by Higgins. Also, the word "significant" should be changed, since readers might interpret that to mean a statistical test, when what you really mean is "substantial heterogeneity. The Cochrane manual says that an I2 of 50%-90% "may represent substantial heterogeneity". Page 7 line 31. Hopefully, for clarity, your sentence can be appropriately changed to "If there were zero events in either group, a value of 0.5 was added to each of the four cells of the 2x2 table". If you only add 0.5 to the zeroes themselves, the resulting effect size will be biased. Page 8 line 31 refers to the EMPA-REG trial as reference 125. This made me think that EMPA-REG was a nonrandomized propensity analysis by Nadkani. However, more sleuthing revealed that in fact EMPA-REG was by Zinman in 2015. Please change the reference numbers so that readeers are not mislead as I was. Also, please add that EMPA-REG was a low risk of bias trial; the fact that it dominated your meta-analyses made me wonder of its risk of bias. Also please state in the results section that ~90% of the patients (or whatever % it is) that were in the meta-analysis of AKI were in that EMPA-REG study.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1 Comments:

C1: The author attempts to address a "knowledge gap" surrounding post-market serious safety outcomes of (SGLT2) inhibitors identified by the FDA and EMA. To do this they conducted a systematic review and meta-analysis of randomized controlled trials (RCT) and used random effects models to estimate pooled relative risks. The primary outcomes were: Acute kidney injury (AKI), diabetic ketoacidosis (DKA), urinary tract infections (UTI), bone fractures, and amputations. They found 99/1865 citations meeting their criteria. When compared to placebo, SGLT2 inhibitors were found to be significantly protective against AKI, while no difference was found for DKA, UTI, or bone fracture. No increased risk for either outcome was found when compared to active controls, and no

studies reported on amputations. Sub-group analysis did show an increased risk of UTI with dapagliflozin only, but no other analysis supported an increased risk of AKI, DKA, UTI, or fracture. The authors conclude that current evidence does not suggest an increased risk of harm with SGLT2 inhibitors as a class over placebo or active comparators with respect to the AKI, DKA, UTI or fracture. However, wide confidence intervals for many comparisons suggest limited precision, and therefore clinically important adverse events cannot be ruled out.

This is a thoughtful and well-done study. It has several strengths and some weaknesses. On the positive side the study is comprehensive on a timely subject. However, they covered events reported as having occurred, there is no information regarding the pathophysiology surrounding the events, i.e. volume depletion as a cause of AKI and urinary incontinence with UTIs. Moreover, certain outcomes are inadequately characterized within study reports. So, the major limitation of this analysis is its dependence the description of other trials.

R1: Thank you for your comments. As suggested, we have added, revised, or expanded previous descriptions with respect to the pathophysiology of the events of interests. Specific changes include:
• The discussion (Page 10, Lines 23-31) three possible mechanisms are described for the relationship between SGLT2 inhibitors and AKI
• The discussion (page 11, lines 6-11) A section has been added to highlight the potential mechanisms for the increased risk of DKA
• The discussion (Page 11, Lines 30-32) describes how increased urinary glucose excretion is thought to be the mechanism for increased UTI, however at this point evidence does not support any increased risk. I could not find any evidence to suggest that users of SGLT2 inhibitors have an increased to urinary incontinence. This is common among patients with type 2 diabetes in general, however our whole population has type 2 diabetes, so this increased risk should not play a role here.
• The discussions (Page 12, Fracture line 6-10) highlights that a disruption in calcium-phosphate homeostasis is thought to be a potentially mechanism for increased fracture risk. However, to date, no increased of fracture has been found.

C2: Beyond these comments this reviewer would caution about comments based on limited data, i.e. Dapa being accused of more UTIs than the others-that is based strictly on observation of some studies and is not backed by any pathophysiological reason for this to occur. Hence, it is a “nonstarter”. The authors should consider adding under each of the uncommon untoward effects a short paragraph alluding a possible pathophysiological basis for this certainly exist for all the events noted. That will add credibility to the paper.

R2: With respect to the increased risk of UTI with only dapagliflozin, additional rationale has been provided (Discussion, Page 11, Line 42 – Page 12, Line 3)

Reviewer 2 Comments:

Donnan et al conducted a systematic review and meta-analysis aiming to evaluate the safety of SGLT2 inhibitors. They included 99 RCTs in their analysis and concluded that SGLT2 inhibitors are not associated with increased risk of harm compared with placebo or active comparators with respect to the AKI, DKA, UTI or fracture. The study is well conducted and written. However, some concerns to be addressed include:

C1: The search strategy of the current study was until July 2017. It would be of importance to updated the search to evaluate if more recent studies have been published from July 2017 till now.

R1: The search has been updated to May 2018. From this search 9 additional studies met our inclusion criteria. Results within the manuscript and online appendix have been modified to reflect this update.

C2: The authors mentioned that no data were available regarding the risk of amputations with SGLT-2 inhibitors. However, in a recent study published in New England Journal of Medicine in 2017 (N Engl J Med 2017; 377:644-657), data from CANVAS and CANVAS-R trials showed significant increased risk of amputations (up to 2-fold increase) with Canagliflozin compared to placebo. Furthermore, in a study by Monami et al in Acta Diabetol (2017) 54:411–413, 17 trials other than the CANVAS reported data regarding toe amputations. Authors should review the CANVAS, CANVAS-R, paper by Monami et al, as well as clinicaltrials.gov for these trials and include data comparing amputations with SGLT-2 inhibitors versus other comparators to the current paper.

R2: Data from the CANVAS program (including CANVAS and CANVAS-R) were identified in the updated literature search ((N Engl J Med 2017; 377:644-657) and (Circ 2018; 137(23))). Unfortunately, all data was presented as rates per 1000 years follow-up, and actual numbers of events was not reported in supplementary materials or clinicaltrials.gov. Results reflecting amputations has been added to the results section. We also reviewed the Monami paper, and noted the amputation from NCT01422876 and also reported this in the results.

C3: Authors should report the weighted mean follow-up duration for each outcome reported, with the weight being the size of the population of each trial.

R3: Thank-you for this suggestion. We agree including a weighted mean follow-up time would strengthen this analysis. We however did not feel that this was feasible for this particular study. Mean follow-up time was not reported within the publications for secondary outcomes, such as those evaluated here. Collecting this data would require contacting individual authors, which given the volume of studies, this is not feasible. We also suspect that we would get a fairly low response from authors for such a request as follow up time for individual adverse effects may not even be captured within study data.

C4: Authors reported data regarding UTI, however no data were reported regarding genital infections, which were reported as serious adverse effect of such class in many studies and meta-analyses (Saad et al, Int J Cardiol. 2017;228:352-358). Authors may report the adverse event of genital infections, or at least explain in their discussion the data in the literature about it.

R4: Though UTIs and genital infections are commonly reported in studies together, we decided to not include genital infections as an outcome in the paper. Genital infections have not been identified as an unanticipated post market safety concern. There is also already systematic review and meta-analytic evidence that more clearly identifies the the increased risk of genital infections with SGLT2 inhibitors.

C5: In the introduction, authors mentioned “No meta-analysis on the risk of DKA currently exists.”, however a prior meta-analysis that evaluated the cardiovascular outcomes with SGLT-2 inhibitors evaluated the risk of DKA as well and found no significant increase compared with placebo (Saad et al, Int J Cardiol. 2017;228:352-358). This makes the above sentence incorrect.

R5: Thank you for identifying this error. The sentence has been changed to reflect the evidence generated by Saad et al. (Page 4, Line 35-36)

C6: Authors decided to include different SGLT2 inhibitors as an eligible comparator. This may cause a confounding effect in the reported outcomes. It would be more accurate to limit the comparator arm to placebo or other non- SGLT2 inhibitor antidiabetic medications.

R6: We had planned on including different SGLT2 inhibitors as comparators, in the event that there was sufficient evidence to conduct a network-meta-analysis between individual agents. We actually did not find any studies that compared different SGLT2 inhibitors head-to-head, and no network meta-analyses were reported here. We have removed this from the list of eligible comparators, as it is not applicable to the results presented and may lead to misinterpretation.

C7: For assessment of risk of bias, authors utilized the Cochrane Collaboration domain-based tool for assessing the risk of bias at the level of the included trial. It is recommended to further perform assessment of risk of bias at the level of the reported outcomes using GRADE tool as recommended by the Cochrane book for meta-analysis.

R7: We agree that typically measuring bias at the level of the individual outcomes would be ideal. The Cochrane Collaboration recommends assessing bias for each main outcome or group of outcomes. When planning for this study this issue was discussed and we decided that the group of outcomes we examined, unanticipated adverse effects, fell within the category of “class of outcomes”. We felt that there was minimal difference expected in bias assessment. One difference however that we did consider was under the domain of selective reporting bias. We identified this when supplementing our data through clinicaltrials.gov. We noticed some data from outcomes that were not reported in the published papers. This was most common for ketoacidosis and fracture. To address this difference, we added context around this issue in the limitations section (Page 12 Lines 35-59).

C8: References need to be revised. For example in the discussion the following sentence has a wrong reference “However, a meta-analysis published in 2017, which is the largest to date, included 77 RCTs representing 50,820 patients and found no increased risk of UTIs in SGLT2 inhibitor users (RR 1.05; 95% CI 0.98-1.12).[16]”. I believe the reference should be 14 rather than 16. Please revise the reference list for accuracy.

R8: Thanks for pointing this out, the references had not updated correctly with the referencing software. This has been corrected.

C9: This sentence in the introduction needs language revision “They identified 73 cases of DKA and 19 cases of life-threatening infections that originated as a UTI, had been identified in patients taking a SGLT2 inhibitor.”

R9: The sentence in question has been changed to the following:

“Among patients taking SGLT2 inhibitors, they identified 19 cases of life-threatening infections that originated as a UTI, and 73 cases of DKA.” (Page 4, 31-32)

Reviewer 3 Comments:

C1: When speaking of non-significant differences, the authors should not lump them all together as “not suggesting a difference” or other vague wording. The problem is that some n.s. findings are clearly indicative that there is no important difference (e.g., your analysis of UTI) whereas other n.s. findings are simply inconclusive due to wide CIs (e.g., your analyses of DKA and bone fractures). While it is technically true that most your analyses “do not suggest an increased risk of harm”, many readers will mistakenly interpret that as “evidence shows that there is no increased risk of harm”. To

avoid this reader misinterpretation, you should clearly delineate which n.s. outcomes are inconclusive, and which n.s. outcomes demonstrate that there is no effect (via a narrow CI). For example, consider the data on UTI for specific medications vs active comparators. When you analyzed empagliflozin, the CI was quite narrow (0.85 to 1.19) indicating that there is no difference. By contrast, when you analyzed other SGLT2's, the data are inconclusive (CI 0.31 to 2.15). These should not be presented identically as "no suggestion of a difference" or "no evidence of an effect". Clearly, with empagliflozin, there is enough evidence to assure us of no added UTI risk. The same cannot be said of other SGLT2's.

R1: We agree, not all non-significant findings should be interpreted as no increased risk. We emphasized this point in this conclusion (both abstract and main text). To further address this issue we modified part of the opening paragraph of the discussion to read:

"We found that SGLT2 inhibitors as a class do not appear to increase the risk of DKA, UTI, and bone fracture, and may have a protective effect with respect to AKI, though this effect was heavily weighted by one large RCT."

(Page 10, Lines 11-14)

C2: Page 6 line 19. Change "primary" to "only".

R2: Changed to "The outcomes of this study include..."

C3: Page 7 line 26. You conducted multiple meta-analyses, not just one.

R3: Changed to "We conducted a series of pair-wise random effects meta-analyses…."

C4: Page 7 line 26. Please specify whether the random effects meta-analysis was the Dersimonian and Laird approach, and if so, cite that paper.

R4: The restricted maximum likelihood (REML) method was used. This has been added and cited.

C5: Page 7 line 29. If a trial compared two SGLT2 medications, should it really be considered together with a trial that compared one SGLT2 medication to a non-SFLT2 medication? The former does not measure the added risk of using an SGLT2, but the latter does.

R5: The eligibility criteria to include other SGLT2 inhibitors as comparators was removed. See response to reviewers 2, comment #6.

C6: Page 7 line 32. Tau is a more direct measure of heterogeneity than I2. See the article by Rucker in BMC Med Res Methodol 8(1) p79. The problem with I2 is that it depends on the Ns. Try it yourself: take a meta-analysis and look at I2. Now triple the Ns, but keep all the effect sizes the same. See how I2 increased? Yeah. That's not good. Note that tau would not have increased, since it more purely measures the extent to which effect sizes differed.

R6: Tau2 values have been included in the forest plots in addition to I2, however the interpretation of heterogeneity was not impacted since all of the I2 values were already small.

C7: Page 7 line 34. You say you used 75% as a threshold for "significant" heterogeneity. Do you have a reference for that? I know that 50% is often the threshold used, citing the original I2 paper by Higgins. Also, the word "significant" should be changed, since readers might interpret that to mean a

statistical test, when what you really mean is “substantial heterogeneity. The Cochrane manual says that an I2 of 50%-90% “may represent substantial heterogeneity”.

R7: The threshold was changed to 50% to reflect the more commonly used threshold. This change does not impact the interpretation of our results in any way. The Higgins references was also added. (Page 7, Lines 20-22)

C8: Page 7 line 31. Hopefully, for clarity, your sentence can be appropriately changed to “If there were zero events in either group, a value of 0.5 was added to each of the four cells of the 2x2 table”. If you only add 0.5 to the zeroes themselves, the resulting effect size will be biased.

R8:Thank for picking up on this unclear point. You are correct, 0.5 was added to each cell. The sentence has been changed to: “If there were zero events reported, a default value of 0.5 was added to all groups within that study.” (Page 7, Lines 19-20)

R8: Page 8 line 31 refers to the EMPA-REG trial as reference 125. This made me think that EMPA-REG was a nonrandomized propensity analysis by Nadkani. However, more sleuthing revealed that in fact EMPA-REG was by Zinman in 2015. Please change the reference numbers so that readeers are not mislead as I was. Also, please add that EMPA-REG was a low risk of bias trial; the fact that it dominated your meta-analyses made me wonder of its risk of bias. Also please state in the results section that ~90% of the patients (or whatever % it is) that were in the meta-analysis of AKI were in that EMPA-REG study.

R8: There was a problem with the referencing software updating the bibliography. This problem has been addressed.

VERSION 2 – REVIEW

REVIEWER	George Bakris University of Chicago Medicine
REVIEW RETURNED	19-Jun-2018

GENERAL COMMENTS	none

REVIEWER	Marwan Saad, MD PhD Department of Cardiovascular Medicine, University of Arkansas for Medical Sciences
REVIEW RETURNED	04-Jul-2018

GENERAL COMMENTS	The authors have performed a comprehensive revision of the manuscript, and addressed the comments appropriately.