Dose response of sodium glucose cotransporter-2 inhibitors in relation to urinary tract infections: a systematic review and network meta-analysis of randomized controlled trials

Jennifer R. Donnan BScPharm MSc, Catherine A. Grandy, Eugene Chibrikov PhD, Carlo Marra PharmD PhD, Kris Aubrey-Bassler MD, Karissa Johnston PhD, Michelle Swab MLIS, Jenna Hache BScPharm, Daniel Curnew BScPharm, Hai Nguyen PhD, John-Michael Gamble BScPharm PhD

Background: The sodium glucose cotransporter-2 (SGLT2) inhibitors are a novel group of drugs for treatment of type 2 diabetes mellitus. We investigated whether there is a dose–response relation between SGLT2 inhibitors and urinary tract infections (UTIs) in patients with type 2 diabetes, relative to other diabetes therapies or placebo.

Methods: We conducted a systematic review and network meta-analysis of randomized controlled trials (RCTs) of SGLT2 inhibitors in patients with type 2 diabetes. We searched 6 databases and the reference lists of key papers. We included studies with placebo or active antidiabetic comparators that reported the outcome of UTI, and established thresholds for high and low doses of SGLT2 inhibitors. We used a random-effects model to estimate the pooled effect estimates and 95% credible intervals.

Results: We screened 2418 citations and included 105 references for studies of 8 unique SGLT2 inhibitors, representing 60 082 individuals (with a total of 4348 UTIs). Most mixed-treatment comparisons showed no significant difference in risk of UTI, with the exception of high-dose dapagliflozin (≥10 mg) compared with placebo (odds ratio [OR] 1.30, 95% credible interval 1.09–1.57), with active comparators (OR 1.44, 95% credible interval 1.04–1.80) and high (OR 1.39, 95% credible interval 1.12–1.72) and low-dose ertugliflozin (OR 1.43, 95% credible interval 1.01–2.01). When the analysis was restricted to RCTs with a low risk of bias, the results were nonsignificant.

Interpretation: Current RCT evidence does not suggest a dose–response relation between most SGLT2 inhibitors and UTIs, with the exception of dapagliflozin. Further research is needed to quantify the relation between SGLT2 inhibitors and more serious infections. Trial registration: PROSPERO registration no. CRD42016038715.

The sodium glucose cotransporter-2 (SGLT2) inhibitors are a novel group of drugs for the treatment of type 2 diabetes mellitus. These products have several benefits, including a moderate glycemic-lowering effect, low risk of hypoglycemia, reductions in weight and blood pressure, and reduction in major adverse cardiovascular events.1,2 The SGLT2 inhibitors are recommended as one of several options for second-line therapy, with empagliflozin and canagliflozin specifically recommended in clinical guidelines as the preferred second-line therapies for patients with pre-existing cardiovascular disease.3 Their unique mechanism of action — inhibition of the reabsorption of glucose at the proximal renal tubule — results in increased urinary glucose excretion4 and has led to speculation about an increased risk of urinary tract infections (UTIs).5 According to a public safety advisory in the United States, there were 19 reported cases of life-threatening kidney or blood infection between March 2013 and October 2014 that originated as a UTI in individuals taking SGLT2 inhibitors.6 Although product monographs for SGLT2 inhibitors identify the increased risk of UTI as a potential adverse effect, clinical trial evidence to date does not support this notion. Two published meta-analyses of randomized controlled trials (RCTs) found no increased risk of UTIs,7,8 except within a subgroup of dapagliflozin users receiving a 10-mg dose,8 which indicated a potential dose–response

Competing interests: None declared.
This article has been peer reviewed.
Correspondence to: John-Michael Gamble, jm.gamble@uwaterloo.ca
CMAJ Open 2018. DOI:10.9778/cmaj.o.20180111
A dose–response relation is plausible, given variation in the rate of urinary glucose excretion with individual agents (Table 1). Some of these agents have shown a clear dose–response relation, whereas others seem to reach a maximum for urinary glucose excretion with certain dosages. Moreover, prior meta-analyses were limited to studies with at least 24 weeks of follow-up. It is unlikely that development of a UTI would require months of treatment, and therefore data from short-term studies should also be considered.

The specific question that we addressed in this systematic review and network meta-analysis was whether there is a

<table>
<thead>
<tr>
<th>Study</th>
<th>Dose (mean 24-h urinary glucose excretion)</th>
<th>Dose response</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Empagliflozin</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhao et al.</td>
<td>Day 1: 10 mg (88 g), 25 mg (83 g)</td>
<td>No difference between 10-mg and 25-mg doses</td>
</tr>
<tr>
<td>Kanada et al.</td>
<td>Day 1: 1 mg (40 g), 5 mg (80 g), 10 mg (85 g), 25 mg (90 g)</td>
<td>Dose response</td>
</tr>
<tr>
<td>Day 27: 1 mg (41 g), 5 mg (77 g), 10 mg (81 g), 25 mg (93 g) (estimated from chart)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scheen</td>
<td>Day 1: 0.5 mg (5 g), 2.5 mg (30 g), 10 mg (50 g), 25 mg (58 g), 50 mg (64 g), 100 mg (80 g), 200 mg (69 g), 400 mg (90.8 g), 800 mg (62 g) (estimated from chart)</td>
<td>Dose response up to about 100-mg dose</td>
</tr>
<tr>
<td>Product monograph</td>
<td>10 mg (64 g), 25 mg (78 g)</td>
<td>Dose response</td>
</tr>
<tr>
<td>Heise et al.</td>
<td>Day 1: 10 mg (74 g), 25 mg (90 g), 100 mg (81 g)</td>
<td>Dose response up to 25-mg dose</td>
</tr>
</tbody>
</table>

| **Dapagliflozin** |  |  |
| Parkinson et al. | 2.5 mg (37.9 g), 5 mg (45.2 g), 10 mg (86.4 g) | Dose response |
| Yang et al. | Day 10: 5 mg (28 g), 10 mg (41 g) | Dose response |
| Product monograph | 10 mg (−70 g); urinary glucose excretion approached a maximum at 20-mg dose | Dose response up to 20-mg dose |

| **Canagliflozin** |  |  |
| Iijima et al. | Ranged from 80 to 110 g; smallest at 25-mg dose, no great difference at 100- to 400-mg dose | Dose response up to 100-mg dose |
| Devineni and Polidori | 100–300 mg (ranged from 80 to 120 g) | Dose response |
| Product monograph | 100–300 mg (ranged from 77 to 119 g) | Not clear |
| Devineni et al. | 50 mg, 100 mg, 300 mg (increased in a dose-dependent manner) | Dose response |

| **Remogliflozin** |  |  |
| Kapur et al. | 20 mg (67 mmol), 50 mg (97 mmol), 150 mg (168 mmol), 500 mg (223 mmol), 1000 mg (304 mmol) | Dose response |
| Dobbins et al. | 200 mg (509 mmol), 1000 mg (918 mmol), 2000 mg (574 mmol) | Dose response up to 1000-mg dose |

| **Ipragliflozin** |  |  |
| Veltcamp et al. | Dose response was noted up to the 50- or 100-mg dose; actual change in urinary glucose excretion depended on the study (50 g in one study and 80–90 g in another study) | Dose response up to 50-mg dose |
| Kadokura et al. | 50 mg (80.6 g ± 22.2 g), 100 mg (89.7 ± 12.3 g) | No difference between 50-mg and 100-mg dose |

| **Ertugliflozin** |  |  |
| Amin et al. | 1 mg (46.33 g), 5 mg (64.54 g), 25 mg (74.49 g) | Dose response |

| **Tofogliflozin** |  |  |
| Ikeda et al. | 2.5 mg (217.9 mmol), 5 mg (272.3 mmol), 10 mg (346.2 mmol), 20 mg (396.0 mmol), 40 mg (402.9 mmol) | Dose response |

| **Sotogliflozin** |  |  |
| Zambrowicz et al. | 400 mg (29.7 g) | Not clear |
| Rosenstock et al. | 75 mg (−18 g), 200 mg (−66 g), 400 mg (55–60 g) (estimated from chart) | Dose response up to 200-mg dose |

Note: SD = standard deviation, SGLT2 = sodium glucose cotransporter-2.
dose–response relation between SGLT2 inhibitors and UTI in individuals with type 2 diabetes, relative to other diabetes therapies or placebo.

Methods

Study design

This study was designed in accordance with the PRISMA statement on systematic reviews and network meta-analyses and was registered with PROSPERO (www.crd.york.ac.uk/prospero; no. CRD42016038715).

Eligibility criteria

For this review, we sought to identify RCTs that compared an SGLT2 inhibitor with placebo, with no treatment or with an active antidiabetic control. The SGLT2 inhibitor could be any one of the currently marketed or investigational agents, but we excluded combined SGLT1/SGLT2 inhibitors. An active control could be any of the available oral antidiabetic agents, with the exception of first-generation sulphonylureas, because they are rarely used in practice. Patients had to be adults (≥ 18 yr) with type 2 diabetes. The studies had to report on the outcome of UTI, but were not limited by duration of follow-up, year of publication or publication status. Inclusion was limited to studies published in English.

Search strategy

A health science librarian (M.S.) conducted a comprehensive literature search. The search strategy was developed in the PubMed database (from inception to May 2018) and was then translated for the Cochrane Library via Wiley (from inception to May 2018), Embase via Embase.com (from inception to May 2018) and International Pharmaceutical Abstracts databases via Ebsco (from inception to May 2018). Medical Subject Headings and keyword terms used to capture type 2 diabetes (e.g., “Diabetes Mellitus, Type 2”[Mesh] OR NIDDM[tw] OR t2dm[tw]) were combined with terms relating to SGLT2 inhibitors, including generic names, brand names, chemical names and compound codes as applicable. RCTs were identified with a methodologic search filter.27 The librarian also conducted multiple test searches to optimize the sensitivity and specificity of the search parameters. Reference lists of key articles were also screened (by J.R.D.). We identified unpublished (grey literature) RCT data by searching the ProQuest Dissertations & Theses Global and ClinicalTrials.gov databases. For the various search strategies, see Section 1 of Appendix 1 (available at www.cmajopen.ca/content/6/4/E594/suppl/DC1).

Study selection and data extraction

DistillerSR software was used to facilitate a 2-level screening process, first with titles and abstracts and then full text (performed by J.R.D., C.A.G., J.H. and D.C.). We used the “liberal accelerated” method of duplicate screening, whereby a second reviewer screens only citations that have been rejected by the first reviewer.

For articles included in the review, one reviewer completed the data extraction, and another performed verification (performed by J.R.D., C.A.G. and J.H.) (for data extraction variables, see Section 2 of Appendix 1). Where gaps existed, the extracted data were supplemented with data from ClinicalTrials.gov. Where data from multiple sources conflicted, information from the published paper was used. Where multiple publications for the same study population existed (e.g., interim analyses or extension studies), the most recent publication was used, except where the most recent publication involved a change in the drug dose.

Assessment of risk of bias

We used the Cochrane Collaboration domain-based risk assessment tool to identify sources of bias in each study.27 This assessment was completed independently by one reviewer, with verification by a second reviewer (performed by J.R.D., C.A.G., J.H.). Each domain was identified as having “low,” “high” or “unclear” risk of bias. In addition, the following rules were applied to assign an overall risk of bias: where all domains were considered to have low risk, the overall risk was low; where at least 1 domain was considered to have high risk, the overall risk was high; and where at least 1 domain was considered to have unclear risk (and no domain was considered to have high risk), the overall risk was considered to be unclear. We assessed publication bias using a funnel plot of placebo-controlled trials.27

Data synthesis

We conducted a Bayesian network meta-analysis of RCTs. The doses of SGLT2 inhibitors were categorized into 2 groups: “high dose” and “low dose.” These categories were defined on the basis of available marketed doses and urinary glucose excretion rates. Where 2 marketed doses were available for a given drug, the lower dose was categorized as “low” and the higher dose as “high.” All other studied doses to the extremes of the 2 marketed doses were categorized in the most proximal dose category. For example, a dose lower than the marketed dose would be considered “low dose.” Where 3 marketed doses were available, the middle dose was categorized with the group having the closest urinary glucose excretion rate. For experimental products, categories were defined by looking at the most commonly studied doses and setting a threshold, as was done for the marketed products. We took this approach to avoid placing too much emphasis on ineffective or unsafe doses used in dose-finding studies (for threshold doses, see Section 3 of Appendix 1).

We used a random-effects generalized linear model for binary data, with non-informative priors, to estimate the relative effects, credible intervals and rank probabilities of each of the comparators. We tested convergence of the Markov Chain Monte Carlo simulation (100 000 iterations) with the Gelman–Rubin diagnostic test and used the deviance information criterion to assess model fit. We examined rank probabilities by calculating the surface under the cumulative rank curve values. We tested heterogeneity with the $F$ statistic for pairwise comparisons and assessed inconsistency by visually comparing the direct and indirect pooled estimates.
We conducted 3 sensitivity analyses. We altered the threshold between low and high doses to reflect uncertainty in the dose–response relation with urinary glucose excretion. We also restricted the analysis to studies of at least 24 weeks’ duration. Finally, we restricted the analysis to studies with a low overall risk of bias.

All of the outcome data were analyzed using the gemtc package of R statistical software (version 3.4.1).

**Ethics approval**
This study was a retrospective analysis of previously published data, and ethics approval was not required.

**Results**
In total, 2418 titles and abstracts were screened, and 140 citations met our inclusion criteria. Of these, 35 were excluded because they represented duplicate data (extension studies, post hoc analyses) or because mixed doses or unstable doses were used. A final list of 105 publications was included in the analysis, representing 108 randomized populations (Figure 1), 60 082 individuals and 4348 UTIs. Three of the publications reported on more than 1 randomized population (for the reference list of included studies, see Section 4 of Appendix 1).

Most studies examined either dapagliflozin (33 studies), empagliflozin (25 studies), canagliflozin (19 studies) or ipragliflozin (11 studies); 20 studies investigated 1 of 4 other agents (luseogliflozin, remogliflozin, tofogliflozin and ertugliflozin). With respect to comparisons, 4 studies conducted only within-class comparisons, 89 compared the study drug with placebo, 26 used an active comparator, and 9 included more than 1 unique comparator. Studies ranged from 1 to 208 weeks in duration (for complete study characteristics, see Section 5 of Appendix 1).

We included all of the studies in the first run of our analysis. However, despite 200 000 iterations of the Markov Chain Monte Carlo simulation, assessment of the Gelman–Rubin

**Figure 1:** Flow diagram for included studies. Initial database searches were from inception of each particular database to May 2018.
statistic showed that many nodes did not approach convergence. There were also unexpected protective effects in comparisons that included luseogliflozin. On examination of study results, we found that only 2 cases of UTI were reported across the 4 luseogliflozin studies, each of which was of short duration (7 d–24 wk). After removal of these studies, all nodes approached convergence. The deviance information criterion was also lower, indicating a better model fit. Figure 2 shows the network of available direct evidence without luseogliflozin.

Most comparisons showed a nonsignificant difference in the risk of UTI (Table 2). Exceptions included comparisons of high-dose dapagliflozin (≥ 10 mg) with placebo (odds ratio [OR] 1.30, 95% credible interval 1.09–1.57), with active comparators (OR 1.44, 95% credible interval 1.15–1.79), with empagliflozin at both high doses (OR 1.39, 95% credible interval 1.12–1.72) and low doses (OR 1.30, 95% credible interval 1.04–1.60) and with ertugliflozin at low doses (OR 1.43, 95% credible interval 1.01–2.01). Low-dose canagliflozin compared with active comparators also had significantly greater risk (OR 1.29, 95% credible interval 1.03–1.64). Examination of rank probabilities using surface under the cumulative rank curve values showed results that were consistent with the primary analysis. Specifically, high-dose dapagliflozin was

---

**Figure 2:** Network diagram for risk of urinary tract infection with sodium glucose cotransporter-2 inhibitors. Note: cana = canagliflozin, dapa = dapagliflozin, empa = empagliflozin, ertu = ertugliflozin, ipra = ipragliflozin, remo = remogliflozin, tofo = tofogliflozin. For each drug, low = low dose, and high = high dose.
the least favourable and high-dose renoglizofin and active comparators (grouped) were the most favourable with respect to risk of UTI (for the forest plot of placebo treatment comparisons and the list of surface under the cumulative rank curve values, see Section 6 of Appendix 1).

Examination of the F value for each of the comparisons showed homogeneity, with most values of F at 0% (and all < 45%). When we back-calculated indirect risk estimates and compared them with direct evidence to assess for consistency, we found no major discrepancies between the estimates, which suggested that the consistency assumption was met (for the complete list of pairwise indirect and pooled estimates, see Section 7 of Appendix 1).

### Table 2: Risk of urinary tract infection in association with SGLT2 inhibitor therapy, as reported for network meta-analysis comparisons*

<table>
<thead>
<tr>
<th>Drug and dose</th>
<th>odds ratio (95% credible interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td>placebo</td>
<td>1.00 (0.91–1.08)</td>
</tr>
<tr>
<td>remo - high</td>
<td>1.07 (0.98–1.17)</td>
</tr>
<tr>
<td>dapa - low</td>
<td>1.03 (0.94–1.13)</td>
</tr>
<tr>
<td>ipragliflozin</td>
<td>1.00 (0.91–1.10)</td>
</tr>
<tr>
<td>low</td>
<td>1.00 (0.91–1.10)</td>
</tr>
<tr>
<td>empagliflozin</td>
<td>1.00 (0.91–1.10)</td>
</tr>
<tr>
<td>remo - low</td>
<td>1.00 (0.91–1.10)</td>
</tr>
<tr>
<td>empagliflozin</td>
<td>1.00 (0.91–1.10)</td>
</tr>
<tr>
<td>placebo</td>
<td>1.00 (0.91–1.10)</td>
</tr>
<tr>
<td>low</td>
<td>1.00 (0.91–1.10)</td>
</tr>
<tr>
<td>empagliflozin</td>
<td>1.00 (0.91–1.10)</td>
</tr>
<tr>
<td>remo - low</td>
<td>1.00 (0.91–1.10)</td>
</tr>
<tr>
<td>low</td>
<td>1.00 (0.91–1.10)</td>
</tr>
</tbody>
</table>

**Note:** SGLT2 = sodium glucose cotransporter 2; (where cana = canagliflozin, dapa = dapagliflozin, emp = empagliflozin, erlu = eruditigliflozin, ipra = ipragliflozin, remo = remogliflozin, tolbo = tolbutamid, and for each, low = dose, and high = high dose).

*Cells containing statistically significant results are indicated with shading.

**Risk of bias**

Generally, the studies were of high methodologic quality. The overall quality assessment indicated that more than half of the studies (54 or 51%) were at low risk of bias. About one-third (32 or 30%) had unclear reporting of randomization sequence, and one-quarter (26 or 25%) had unclear risk of bias for blinded outcome assessment (Figure 3). No indication of publication bias was observed in the funnel plot (Section 8 of Appendix 1).

**Sensitivity analysis**

The results of the sensitivity analyses were consistent with those of the primary analysis. When the threshold between
high and low doses was altered, high-dose dapagliflozin still showed an increased risk of UTI compared with placebo, active comparators and high-dose empagliflozin, but also showed an increased risk relative to low doses of ipragliflozin and ertugliflozin. The thresholds for dapagliflozin doses were not adjusted in this sensitivity analysis, because an alternative definition was not suitable. Ipragliflozin at low doses showed a significantly lower risk of UTI than high doses of canagliflozin, ertugliflozin, ipragliflozin and dapagliflozin.

In the analysis of studies lasting 24 weeks or longer, fewer comparisons among experimental agents were possible. However, the findings were consistent with those of the primary analysis, whereby high-dose dapagliflozin had a high risk compared with placebo, active comparator and empagliflozin. Restriction of the analysis to studies with an overall low risk of bias (n = 57) resulted in no significant differences among the drug regimens.

In each of the sensitivity analyses, there were treatment arms with insufficient data to accurately estimate risk (for complete results of the sensitivity analyses, see Section 9 of Appendix 1).

**Interpretation**

The main findings of this study suggest no dose–response association between SGLT2 inhibitors and UTI risk; however, dapagliflozin (at doses ≥ 10 mg) appears to be an exception to this general finding. Specifically, high-dose dapagliflozin compared with placebo, active comparator and empagliflozin was associated with a small increase in the risk of UTI.

Several other meta-analyses have reported on the association between SGLT2 inhibitors and UTIs, with inconsistent results, including increased risk with dapagliflozin, increased risk with SGLT2 inhibitors, and no difference in risk. However, given the continuing postmarketing surveillance of these new agents, new RCTs are being published rapidly, and these previous meta-analyses are quickly becoming outdated. In addition, several studies have applied additional eligibility criteria, such as including only marketed agents, placebo comparison trials or studies of a certain duration (e.g., > 24 wk). The largest meta-analysis to date, which pooled results from 86 RCTs representing 50 880 patients, found no increased risk of UTIs (relative risk 1.03, 95% confidence interval 0.96–1.11). Subgroup analysis by dose in this previous study also showed an increased risk only among users of dapagliflozin at a 10-mg dose.

A mechanism for the increased risk of UTI with dapagliflozin is not clear; however, there is variation in the pharmacokinetic and pharmacodynamic profiles of individual SGLT2 inhibitors. The SGLT2 inhibitors have shown a positive dose–effect relation with urinary glucose excretion, but this appears to have a ceiling effect with several agents. Maximum effects have been documented at about the starting doses for empagliflozin (10 mg) and canagliflozin (100 mg), but continued through the dosing range with dapagliflozin. This may explain why the current study showed a dose-dependent relation for UTIs with dapagliflozin. It is unclear why an increased risk of UTI was observed with low-dose canagliflozin. Our sensitivity analysis showed a potential decreased risk of UTI among users of low-dose ipragliflozin and high doses of dapagliflozin; or placebo. Pharmacodynamic evidence for iragliflozin has been variable, with inconsistent estimates of...
the degree of urinal glucose excretion and the dose–response relation. However, there is also no indication that ipragliflozin is unique in any way that would support a physiologic mechanism for the decreased risk of UTI. Further work is needed to examine this finding.

Our findings are consistent with previous findings supporting a lack of compelling data that would suggest a class effect in terms of UTI risk. Our study also extends the evidence by including additional studies, which has resulted in a more precise effect estimate. This study included as many studies as we could find to investigate dose response encompassing both marketed and nonmarketed agents, and active and inactive comparators in studies of any duration.

Limitations
This systematic review of the association between SGLT2 inhibitors and UTIs had some limitations. The outcome of UTI is very well reported, but we did not identify data on the progression of UTI to more serious infections. This gap in reporting makes it impossible to support or refute the concern that SGLT2 inhibitors may lead to serious infections. It is already known that, as a population, patients with diabetes have an increased risk of infections of all origins. The 19 serious cases of UTI associated with SGLT2 inhibitors reported in the United States may be a result of increased vigilance for newly marketed drugs. The role of urinary glucose excretion in the pathogenesis of urinary tract infections is not well characterized. It has been postulated that increased urinary glucose excretion may not directly cause infections but rather may create a rich environment for bacterial growth and affect bacterial adherence to uroepithelial cells. Because of the volume of studies included, it was not feasible to contact authors regarding these data. Other limitations included restriction of the analysis to studies published in English, and verification of data abstraction and bias assessment by a second reviewer, rather than independent duplication of abstraction and assessment. Finally, we found no study that compared 2 different SGLT2 inhibitors in a single trial; therefore, the strength of evidence for comparisons between SGLT2 inhibitors is weak.

Conclusion
Current evidence does not support a dose–response risk profile for UTIs with SGLT2 inhibitors as a class. Although high doses of dapagliflozin (≥ 10 mg) did appear to be associated with increased risk, this risk was attenuated in an analysis restricted to RCTs with low risk of bias. Further studies are needed to quantify the association between SGLT2 inhibitors and more serious infections such as pyelonephritis.

References


**Affiliations:** School of Pharmacy (Donnan, Grandy, Chibrikov, Marra, Johnston, Hache, Curnew, Nguyen, Gamble) and Faculty of Medicine (Aubrey-Bassler, Swab), Memorial University, St. John’s, NL; School of Pharmacy (Marra), University of Otago, Dunedin, New Zealand; School of Pharmacy (Gamble), University of Waterloo, Kitchener, Ont.

**Contributors:** Jennifer Donnan led the review and was involved at every stage, including protocol development, search strategy design, screening, data extraction, quality appraisal, analysis and manuscript preparation. Catherine Grandy was involved in screening, data extraction, quality appraisal and manuscript revisions. Eugene Chibrikov was involved in data cleaning and analysis, and manuscript revisions. Carlo Marra, Kris Aubrey-Bassler and Karissa Johnston were involved in project conception, protocol development and manuscript revisions. Michelle Swab was involved in search strategy design, literature search and manuscript revisions. Jenna Hache and Daniel Curnew were involved in screening, data extraction, quality appraisal and manuscript revisions. Hai Nguyen was involved in interpretation of study results and manuscript revisions. John-Michael Gamble supervised this research and was involved in protocol development, consensus on disagreements in data extraction, data analysis, interpretation of results and manuscript revisions. All of the authors gave approval of the version to be published and agreed to be accountable for all aspects of the work.

**Supplemental information:** For reviewer comments and the original submission of this manuscript, please see www.cmajopen.ca/content/6/4/E594/suppl/DC1.