Safety and effectiveness of dipeptidyl peptidase-4 inhibitors versus intermediate-acting insulin or placebo for patients with type 2 diabetes failing two oral antihyperglycaemic agents: a systematic review and network meta-analysis

Andrea C Tricco,1 Jesmin Antony,1 Paul A Khan,1 Marco Ghassemi,1 Jemila S Hamid,2 Huda Ashoor,1 Erik Blondal,1 Charlene Soobiah,1 Catherine H Yu,1 Brian Hutton,3 Brenda R Hemmelgarn,4 David Moher,3 Sumit R Majumdar,5 Sharon E Straus1,6

ABSTRACT

Objective: To evaluate the effectiveness and safety of dipeptidyl peptidase-4 (DPP-4) inhibitors versus intermediate-acting insulin for adults with type 2 diabetes mellitus (T2DM) and poor glycaemic control despite treatment with two oral agents.

Setting: Studies were multicentre and multinational.

Participants: Ten studies including 2967 patients with T2DM.

Interventions: Studies that examined DPP-4 inhibitors compared with each other, intermediate-acting insulin, no treatment or placebo in patients with T2DM.

Primary and secondary outcome measures: Primary outcome was glycosylated haemoglobin (HbA1c). Secondary outcomes were healthcare utilisation, body weight, fractures, quality of life, microvascular complications, macrovascular complications, all-cause mortality, harms, cost and cost-effectiveness.

Results: 10 randomised clinical trials with 2967 patients were included after screening 5831 titles and abstracts, and 180 full-text articles. DPP-4 inhibitors significantly reduced HbA1c versus placebo in network meta-analysis (NMA; mean difference (MD) −0.62%, 95% CI −0.93% to −0.33%) and meta-analysis (MD −0.61%, 95% CI −0.81% to −0.41%), respectively. Significant differences in HbA1c were not observed for neutral protamine Hagedorn (NPH) insulin versus placebo and DPP-4 inhibitors versus NPH insulin in NMA. In meta-analysis, no significant differences were observed between DPP-4 inhibitors and placebo for severe hypoglycaemia, weight gain, cardiovascular disease, overall harms, treatment-related harms and mortality, although patients receiving DPP-4 inhibitors experienced less infections (relative risk 0.72, 95% CI 0.57 to 0.91).

Conclusions: DPP-4 inhibitors were superior to placebo in reducing HbA1c levels in adults with T2DM taking at least two oral agents. Compared with placebo, no safety signals were detected with DPP-4 inhibitors and there was a reduced risk of infection. There was no significant difference in HbA1c observed between NPH and placebo or NPH and DPP-4 inhibitors.

Trial registration number: PROSPERO # CRD42013003624.

BACKGROUND

Patients newly diagnosed with type 2 diabetes mellitus (T2DM) may initially be treated with...
lifestyle modifications, including healthy diet and physical activity. However, T2DM is a progressive disease and pharmacological interventions in the form of oral antihyperglycaemic agents (eg, metformin, sulfonylurea, thiazolidinedione, glucagon-like peptide-1)\(^1\)\(^2\)\(^3\) are usually required. For some patients, glycaemic control is not achieved until two oral antihyperglycaemic agents are administered (ie, second-line therapy).\(^2\)\(^3\)

Some patients will not achieve sustained diabetes control even after taking two oral antihyperglycaemic agents.\(^4\) These patients may require third-line therapy, such as dipeptidyl peptidase-4 inhibitors (DPP-4 inhibitors) or intermediate-acting insulin (lente, neutral protamine Hagedorn (NPH)). However, the comparative safety, effectiveness and cost of DPP-4 inhibitors versus intermediate-acting insulin or placebo are unclear.

We aimed to determine the comparative safety and effectiveness of DPP-4 inhibitors compared with intermediate-acting insulin or placebo for adults with T2DM whose glycosylated haemoglobin (HbA1c) levels remain elevated despite the administration of two oral agents.

**METHODS**

Our systematic review protocol was drafted using guidance from the Preferred Reporting Items for Systematic reviews and Meta-analyses for Protocols (PRISMA-P).\(^5\)

We revised our protocol to address feedback received by the study team and relevant knowledge users from the British Columbia Ministry of Health and registered it with the international prospective systematic review register (PROSPERO; CRD42013003624). As our full methods have been published previously,\(^6\) we only outlined them briefly here.

**Eligibility criteria**

We included studies examining DPP-4 inhibitors (eg, sitagliptin, vildagliptin, saxagliptin and linagliptin) and intermediate-acting insulin (eg, NPH, lente) compared with each other or placebo for the third-line treatment of adult patients with T2DM and an HbA1c $\geq 6.5\%$. We defined third-line treatment as when the study examined the use of two oral antihyperglycaemic agents among all patients, plus the addition of a DPP-4 inhibitor, intermediate-acting insulin or placebo (ie, three agents in total per group). This is consistent with previous reviews examining third-line treatment of T2DM.\(^7\)\(^9\)

HbA1c was our primary outcome of interest, while healthcare utilisation (eg, emergency department visits), body weight, fractures, quality of life, microvascular complications (retinopathy, neuropathy, nephropathy), macrovascular complications (cardiovascular disease, stroke/transient ischaemic attack, peripheral vascular disease), all-cause mortality, harms (including infection, pancreatic cancer, severe hypoglycaemia, serious hyperglycaemia and body weight), cost and cost-effectiveness were our secondary outcomes. Studies were included regardless of duration of follow-up, date of dissemination/publication, language of dissemination/publication and publication status (ie, unpublished studies were eligible for inclusion). All study designs were included with the exception of qualitative studies, case reports, case series and cross-sectional studies.

**Information sources and literature search**

Published and unpublished studies were identified through searching electronic databases (MEDLINE, EMBASE and the Cochrane Central Register of Controlled Trials), clinical trial registries (eg, WHO International Clinical Trials Search Portal) and drug manufacturer sites from inception until 18 December 2012. This search was supplemented using various methods, including most related articles searching for each of the included studies in PubMed, searching articles that reference the included studies in Web of Science (also called forward citation searching), and scanning the reference lists of included studies and relevant reviews.\(^7\)\(^9\)

The literature search for MEDLINE was drafted by an experienced information specialist and peer reviewed by another using the Peer Review of Electronic Search Strategies (PRESS) checklist.\(^10\) The full MEDLINE search has been published previously,\(^6\) and searches for the other databases are available from the authors on request.

**Study selection process**

After conducting a calibration exercise, each title and abstract from the literature search was reviewed by two team members independently using our synthesi.sr tool.\(^11\) Conflicts were resolved by discussion. The same process was followed for screening potentially relevant full-text articles.

**Data items and data collection process**

Data items included study characteristics (eg, setting, country of conduct, intervention and comparator examined), patient characteristics (eg, duration of diabetes, mean age) and outcome results. After a calibration exercise, each of the included studies was abstracted by two team members independently. Conflicts were resolved by discussion. The data abstraction forms are available from the authors on request.

**Risk of bias and methodological quality appraisal**

The included studies were appraised using the seven-item Cochrane risk-of-bias tool.\(^12\) Furthermore, studies reporting harms were assessed using the 15-item McMaster Quality Assessment Scale of Harms (McHarm) tool,\(^13\) which focuses on how harms are defined, collected and reported. Each included study was independently assessed by two members of the review team and any discrepancies were resolved through discussion.

**Synthesis**

Study and patient characteristics and risk of bias/methodological quality results were summarised
When sufficient data were available, random effects meta-analysis was conducted to calculate the pooled mean difference (MD) for continuous outcomes and relative risk (RR) for dichotomous outcomes. Clinical, methodological and statistical (eg, I² statistic) heterogeneity were considered. Subgroup analysis was conducted when significant heterogeneity was observed (eg, I² statistic >75%). These analyses were conducted using R statistical software. Missing data (eg, SE of estimates) were imputed using established methods. The impact of these imputations was examined through sensitivity analysis.

In addition, we conducted a Bayesian random effects network meta-analysis (NMA) using R and WinBUGS for HbA1c, the primary outcome of interest. Median rankings and 95% credible intervals were calculated using all available direct (ie, from head-to-head trials) and indirect (ie, from Bayesian NMA) data. The 95% credible interval is generated by the Bayesian NMA and is interpreted in the same manner as the 95% CI generated by traditional meta-analysis. The surface under the cumulative ranking curve (SUCRA) was used for ranking the treatments. Treatments were grouped into nodes with input from clinicians on the team and the robustness of the selected treatment nodes was examined via sensitivity analysis. Trace and history plots were visualised and the Gelman Rubin statistic was calculated. Consistency of indirect and direct results was examined using established statistical methods. Important network inconsistency was explored using sensitivity analysis. We assessed the transitivity assumption by examining the comparability of the distribution of the treatment effect modifiers across comparisons, including HbA1c levels (<8% vs ≥8%).

RESULTS

Literature search

The literature search yielded a total of 5831 titles and abstracts (figure 1). Of these, 180 full-text articles were potentially relevant. Ten studies fulfilled the eligibility criteria and were included.5–34 Five of the included studies were unpublished,26–28 30 31 four of which were identified as conference abstracts in our literature search,26–28 31 while one was a protocol with results identified through searching the trial registries.30 An unpublished dissertation was found from one of the conference abstracts and translated into English from Portuguese.31 All of the included studies were randomised clinical trials (RCTs).

Study and patient characteristics

The study durations ranged from 12 to 36 weeks and a total of 2967 adults with T2DM were included (table 1). All RCTs examined two oral antihyperglycaemic agents plus a DPP-4 inhibitor, including saxagliptin,25–27 29 31 32 34 linagliptin26 33 and vildagliptin.27 28 Eight RCTs were placebo-controlled; patients in these arms also received the same two oral antihyperglycaemic agents as those in the intervention arm.25–26 28–30 32–34 One RCT compared sitagliptin plus metformin and a sulfonylurea with vildagliptin plus metformin and a sulfonylurea,27 while another compared sitagliptin plus metformin and a sulfonylurea with NPH insulin plus metformin and a sulfonylurea.31 All studies examined metformin plus the addition of a sulfonylurea, pioglitazone or exenatide in all arms. None of the included RCTs examined the intermediate-acting insulin lente.

At baseline, the proportion of women ranged from 37.7% to 57.1% and mean age of participants ranged from 49.8 to 58.1 years (table 2). The mean body mass index ranged from 28.2 to 35.2 kg/m² at baseline, while mean HbA1c ranged from 7.7% to 8.8% at baseline. The average duration of disease ranged from 7.3 to 10.9 years. All of the patients were obese in one RCT,29 while 69% had high cholesterol and 89% had hypertension in another RCT.31 None of the other RCTs reported comorbidities.

Risk of bias and methodological quality results

Only one RCT was at a low risk of bias for random sequence generation.34 two RCTs had a low risk of allocation concealment bias25 34 and all had a low risk of bias due to blinding because the primary outcome was objective (table 3, see online supplementary appendix 1). Five RCTs had a high risk of incomplete outcome data bias as more than 10% of patients withdrew from the trial despite the short duration of the follow-up (median of
24 weeks) or as there was an imbalance in the number of patients withdrawing between the groups.25 26 30 32 34 Five RCTs were at a high risk of ‘other type of bias’ (eg, funding bias) as the trials were funded by a pharmaceutical company and some of the authors were employed by the funding company.25 26 32–34 Four RCTs had an unclear risk of ‘other type of bias’ (eg, funding bias) as some were funded by a pharmaceutical company (none of the authors were employed by the company) or as it was an unpublished study that failed to report the funding source and further information was required.27 28 30 31

Seven of the RCTs reported on harms and were assessed using the McHarm tool13 (see online supplementary appendices 2 and 3). The proportion of items outlined in online supplementary appendix 2 that were adequately reported across the studies was low, with most RCTs reporting 26–53% of the items. However, one RCT adequately reported 73% of the items.34

### NMA results

**Primary outcome—HbA1c**

Patients receiving a DPP-4 inhibitor (plus metformin and a sulfonylurea/exenatide/pioglitazone) had an average reduction in HbA1c (MD 0.62%, 95% credible interval 0.33 to 0.93, I²=87%) versus placebo plus metformin and a sulfonylurea/exenatide/pioglitazone in NMA including eight RCTs (figure 2). Patients receiving NPH insulin experienced a reduction in HbA1c compared to placebo (MD 0.62%, 95% credible interval 0.31 to 1.57); however, this reduction was not statistically significant, which may be attributed to insufficient power due to the limited number of studies providing data on these treatment comparisons. No statistically significant difference was observed between DPP-4 inhibitors and NPH insulin (MD 0.00%, 95% credible interval −0.89 to 0.89). However, the ranking and cumulative ranking probabilities indicated that DPP-4 was superior to NPH insulin where the SUCRA values were 74.89%, 70.89% and 4.22% for DPP-4, NPH insulin and placebo, respectively (table 4). Consistency was observed between direct and indirect evidence visually and statistically using the node splitting approach for the closed loop.

Because of the statistically significant heterogeneity identified in the NMA, we conducted several unplanned sensitivity analyses. A sensitivity analysis was conducted with each of the drugs and these were included separately in the model (ie, without grouping the DPP-4 inhibitors together). Only sitagliptin plus metformin and a sulfonylurea, and vildagliptin plus metformin and a sulfonylurea significantly reduced HbA1c versus placebo plus metformin and a sulfonylurea (MD −0.83%, 95% credible interval −1.55 to −0.14 and MD −0.97%, 95% credible interval −1.89 to −0.19, respectively). According to SUCRA values (table 5), vildagliptin plus metformin and a sulfonylurea had the greatest probability (85%) of being the most effective in reducing HbA1c compared with all agents included in the model, while
sitagliptin plus metformin and a sulfonylurea had the second greatest probability (76%) of being the most effective in reducing HbA1c. Furthermore, sensitivity analyses were conducted to control for baseline HbA1c values and account for imputations on missing data; no changes in the results were observed.

Meta-analysis results

Primary outcome—HbA1c

After a median of 24 weeks of follow-up, patients receiving a DPP-4 inhibitor plus metformin and a sulfonylurea/exenatide experienced a significantly reduced HbA1c compared with those receiving placebo plus metformin and a sulfonylurea/exenatide (5 RCTs, MD −0.61%, 95% CI −0.81% to −0.41%, I²=87%).

Secondary outcome—microvascular complications

A meta-analysis was not conducted for microvascular complications because only one RCT reported neuropathy. In this study, no differences were observed for patients receiving saxagliptin plus metformin and a sulfonylurea versus placebo plus metformin and a sulfonylurea after 24 weeks of follow-up (RR 6.95, 95% CI 0.36 to 133.13). None of the included RCTs reported retinopathy or nephropathy.

Secondary outcome—macrovascular complications

After a median of 22 weeks follow-up, no differences were observed for cardiovascular disease (unspecified or acute myocardial infarction) among patients receiving a DPP-4 inhibitor plus metformin and a sulfonylurea/exenatide versus placebo plus metformin and a sulfonylurea/exenatide (2 RCTs, RR 0.18, 95% CI 0.02 to 1.63, I²=0%). One RCT was not included in the cardiovascular disease meta-analysis as it reported 0 patients experiencing chronic heart failure in both groups. None of the included RCTs reported stroke/transient ischaemic attack or peripheral vascular disease.

### Table 2 Patient characteristics

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Female (%)</th>
<th>Mean age (SD), years</th>
<th>Mean BMI (SD)</th>
<th>Mean HbA1c (SD)</th>
<th>Mean duration of T2DM (SD), years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fonseca et al, 2013</td>
<td>37.7</td>
<td>56.1 (9.06)</td>
<td>30.0 (5.20)</td>
<td>8.8 (1.00)</td>
<td>9.8 (5.95)</td>
</tr>
<tr>
<td>Gilman, 2013*</td>
<td>51.5</td>
<td>53.8 (9.30)</td>
<td>28.2 (5.30)</td>
<td>NR (NR)</td>
<td>NR (NR)</td>
</tr>
<tr>
<td>Abdulwahid, 2012*</td>
<td>45.5</td>
<td>49.8 (12.58)</td>
<td>NR (NR)</td>
<td>8.62 (0.20)</td>
<td>NR (NR)</td>
</tr>
<tr>
<td>Lukashevich, 2012*</td>
<td>NR</td>
<td>55.1 (NR)</td>
<td>28.0 (NR)</td>
<td>8.78 (0.63)</td>
<td>7.3 (NR)</td>
</tr>
<tr>
<td>Makkissi et al, 2012*</td>
<td>45.5</td>
<td>53.5 (12.09)</td>
<td>35.2 (4.74)</td>
<td>7.74 (1.21)</td>
<td>NR (NR)</td>
</tr>
<tr>
<td>Moses et al, 2012*</td>
<td>40.1</td>
<td>57.0 (10.54)</td>
<td>29.2 (5.09)</td>
<td>8.28 (0.85)</td>
<td>NR (NR)</td>
</tr>
<tr>
<td>Nogueira, 2012*</td>
<td>57.1</td>
<td>56.7 (6.80)</td>
<td>27.0 (2.60)</td>
<td>8.1 (0.65)</td>
<td>10.9 (6.68)</td>
</tr>
<tr>
<td>Violante et al, 2012*</td>
<td>49.4</td>
<td>56.0 (7.92)</td>
<td>31.2 (5.19)</td>
<td>7.89 (0.62)</td>
<td>8.0 (6.52)</td>
</tr>
<tr>
<td>Owens et al, 2011*</td>
<td>52.8</td>
<td>58.1 (9.80)</td>
<td>28.3 (4.70)</td>
<td>8.15 (0.04)</td>
<td>NR (NR)</td>
</tr>
<tr>
<td>Hermansen et al, 2007*</td>
<td>47.6</td>
<td>57.2 (8.85)</td>
<td>31.0 (6.05)</td>
<td>8.27 (0.71)</td>
<td>9.9 (6.27)</td>
</tr>
</tbody>
</table>

*Unpublished data.

HbA1c, glycosylated haemoglobin (%); BMI, body mass index (kg/m²); NR, not reported; T2DM, type 2 diabetes mellitus.

### Table 3 Appraisal of risk of bias of the included studies using the Cochrane risk-of-bias tool

<table>
<thead>
<tr>
<th>Study</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fonseca et al</td>
<td>Unclear</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Gilman*</td>
<td>Unclear</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Abdulwahid*</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Low</td>
<td>Low</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Unclear</td>
</tr>
<tr>
<td>Lukashevich*</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Unclear</td>
</tr>
<tr>
<td>Makkissi et al</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Low</td>
<td>Low</td>
<td>Unclear</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Moses et al</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Unclear</td>
<td>Unclear</td>
</tr>
<tr>
<td>Nogueira*</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Unclear</td>
<td>Unclear</td>
</tr>
<tr>
<td>Violante et al</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Owens et al</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Hermansen et al</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Unclear</td>
<td>High</td>
</tr>
</tbody>
</table>

Items:
1. Random sequence generation.
2. Allocation concealment.
3. Blinding of participants and personnel.
5. Incomplete outcome data.
7. Other bias.

*Unpublished data.

High, high risk; Low, low risk; Unclear, unclear risk.
Secondary outcome—all-cause mortality

After a median of 25 weeks follow-up, no differences were observed for all-cause mortality among patients receiving sitagliptin plus metformin and a sulfonylurea/pioglitazone versus placebo plus metformin and a sulfonylurea/pioglitazone (2 RCTs, RR 0.98, 95% CI 0.10 to 9.41, I²=0%).

Secondary outcome—harms

After a median of 24 weeks of follow-up, no differences were observed for any harm (5 RCTs, RR 1.07, 95% CI 0.96 to 1.19, I²=22%, figure 3) or overall harms that were treatment related according to the authors of the trials (4 RCTs, RR 1.38, 95% CI 0.92 to 2.09, I²=67%, figure 3) for patients receiving a DPP-4 inhibitor plus metformin and a sulfonylurea/exenatide/pioglitazone versus placebo plus metformin and a sulfonylurea/exenatide/pioglitazone. After a median of 22 weeks of follow-up, no differences were observed for severe hypoglycaemia (defined as “requiring the assistance of another person to administer carbohydrate, glucagon or resuscitative actions”33 or “any episode with symptoms consistent with hypoglycaemia resulting in loss of consciousness or seizure that showed prompt recovery in response to administration of glucagon or glucose or documented hypoglycaemia [blood glucose <3.0 mmol/l (54 mg/dl)]”, requiring assistance of another person because of severe impairment in consciousness or behavior; whether or not symptoms of hypoglycaemia were detected by the patient35) in patients receiving a DPP-4 inhibitor plus metformin and a sulfonylurea/exenatide versus placebo plus metformin and a sulfonylurea/exenatide (2 RCTs, RR 0.69, 95% CI 0.48 to 1.00).
0.16 to 2.94, I²=0%, figure 3). Two RCTs reporting severe hypoglycaemia were excluded from the meta-analysis as they reported 0 events in both groups.25-34 One RCT32 reported serious hyperglycaemia (undefined) and no differences were observed between sitagliptin plus metformin and exenatide versus placebo after 20 weeks of follow-up (RR 0.33, 95% CI 0.01 to 8.04).

After a median of 24 weeks of follow-up, patients receiving a DPP-4 inhibitor plus metformin and a sulfonylurea/pioglitazone experienced significantly fewer infections overall compared with those receiving placebo plus metformin and a sulfonylurea/pioglitazone (4 RCTs, RR 0.72, 95% CI 0.57 to 0.91, I²=0%, figure 3). A subgroup analysis was conducted by the specific type of infection, and the DPP-4 inhibitors plus metformin and sulfonylurea/pioglitazone/exenatide were not significantly different versus placebo plus metformin and sulfonylurea/pioglitazone/exenatide for upper respiratory tract infection, nasopharyngeal infection and urinary tract infection. None of the studies reported pancreatitis and only one study reported on pancreatic cancer, which was not statistically significant for sitagliptin plus metformin and exenatide versus placebo plus metformin and exenatide (RR 0.33, 95% CI 0.01 to 8.04).32

After a median of 24 weeks of follow-up, patients receiving a DPP-4 inhibitor plus metformin and a sulfonylurea/exenatide/pioglitazone did not experience any differences in body weight compared with those receiving placebo plus metformin and a sulfonylurea/exenatide/pioglitazone (4 RCTs, MD 0.23 kg, 95% CI −1.58 to 2.04, I²=0%). No significant differences were observed for sitagliptin plus metformin and a sulfonylurea versus NPH insulin plus metformin and a sulfonylurea regarding body weight in one RCT31 (MD −4.10 kg, 95% CI −11.32 to 3.12).

**DISCUSSION**

We found a statistically significant and clinically important improvement in HbA1c for DPP-4 inhibitors versus placebo. Although no statistically significant differences were observed between DPP-4 inhibitors and NPH insulin, ranking probabilities indicate that DPP-4 inhibitor has a greater likelihood of reducing HbA1c versus NPH insulin. Clinically important improvement in HbA1c was also observed for NPH insulin versus placebo. However, this difference was not statistically significant and this was most likely due to insufficient power due to the limited number of studies contributing data to this analysis. Furthermore, our NMA results did not change when we controlled for baseline HbA1c values. No other statistically significant differences were observed across all outcomes, except that patients receiving a DPP-4 inhibitor experienced fewer infections overall compared with placebo.

**Figure 3** Harms meta-analysis results. This is the forest plot for all harms meta-analyses. a, Unpublished data. N, sample size.
Regarding the generalisability of our results, one RCT which included obese patients with T2DM compared sitagliptin plus metformin and a sulfonylurea with placebo plus metformin and a sulfonylurea. Similar results were observed in this trial compared with the other RCTs examining DPP-4 inhibitors versus placebo. The only head-to-head trial examining DPP-4 inhibitors versus insulin included a high proportion of patients with T2DM with high cholesterol and hypertension. We are unable to comment on whether these results are representative of patients with T2DM without these characteristics because these comorbidities were not reported in the other included trials. The median baseline HbA1c level was 8.27% (range 7.74–8.80%) across the RCTs, suggesting that the included RCTs are representative of patients with T2DM with inadequate glycaemic control, despite the administration of two oral antihyperglycaemic agents.

Our results are similar to previous reviews on this topic, all of which found significantly reduced HbA1c levels for DPP-4 inhibitors compared with placebo. However, these reviews also observed significantly increased weight gain for DPP-4 inhibitors versus placebo, while we did not. The first two reviews did not examine infections, yet a Cochrane review found significantly increased infections for sitagliptin but not for vildagliptin. This Cochrane review examined DPP-4 inhibitors for any patient with T2DM, whereas we focused on those failing two oral antihyperglycaemic agents. Although the Cochrane review included many more studies examining DPP-4 inhibitors, we included four additional RCTs that were published after the literature search date of 2008 for the Cochrane review, as well as four unpublished RCTs. These eight RCTs were also not included in the other two reviews that had later search dates. None of the previous reviews provided results specifically for DPP-4 inhibitors versus intermediate-acting insulin, such as NPH.

Our results are also similar to two recent randomised trials that examined a DPP-4 inhibitor versus placebo among patients with T2DM and acute coronary syndrome or at risk for cardiovascular events. One trial found that no differences were observed between a DPP-4 inhibitor and placebo regarding major cardiovascular events after 40 months follow-up. The other trial found that after 2.1 years follow-up, DPP-4 inhibition was not significantly different than placebo for ischaemic events, yet the risk of hospitalisation for heart failure was increased for saxagliptin versus placebo. We were unable to include these trials in our systematic review as they did not include only patients with T2DM failing two oral antihyperglycaemic agents. Furthermore, neither of these trials examined the effects of intermediate-acting insulin versus DPP-4 inhibitors.

When planning this review, we aimed to include studies providing ‘real-world’ data comparing these agents, such as cohort studies. However, we did not identify any such study through our extensive literature search that fulfilled our eligibility criteria. We also did not identify studies examining other important outcomes of interest, including healthcare utilisation, fractures, quality of life, cost and cost-effectiveness. These are important outcomes for consideration in future trials.

The studies included in our review can be improved by adequately reporting harms (only one of the RCTs adequately reported the majority of the harms on the McHarm tool), as well as reporting patient characteristics to allow clinicians’ assessment of external validity of the results. Furthermore, many of the trials had a short duration of follow-up (the longest duration was only 36 weeks), suggesting that studies of longer duration are required in this area to assess for durability of glycaemic control as well as impact on development of long-term complications and mortality. This issue is especially critical given that T2DM is a chronic disease and therapies are implemented long term.

Our NMA of HbA1c should be interpreted with caution, as only eight RCTs were included in this analysis. As such, we did not formally test for publication bias but assume that our results are not affected by this, as 50% of our included studies were unpublished trials. Although we sought to include non-English studies, only one was identified, suggesting that we may have missed studies written in languages other than English.

In conclusion, our results suggest that DPP-4 inhibitors are superior to placebo and have similar effectiveness as NPH insulin in reducing HbA1c as a third-line therapy in patients with T2DM. Patients and their healthcare providers may rely on patient preferences and other factors when selecting a DPP-4 inhibitor versus insulin for patients with T2DM inadequately controlled by two oral antihyperglycaemic agents. Our results are also relevant to decision-makers, such as the British Columbia Ministry of Health, who posed this query because they required effectiveness and safety data for DPP-4 inhibitors versus intermediate-acting insulin. Unfortunately, we did not identify any study comparing the costs of these agents and so are unable to provide any information on this important factor. This literature base can be improved by ensuring less patient dropouts, adequate reporting of patient characteristics and harms, and examining important diabetes outcomes, including healthcare utilisation, fractures, quality of life, cost and cost-effectiveness.

Author affiliations
1Li Ka Shing Knowledge Institute, St. Michael’s Hospital, Toronto, Ontario, Canada
2Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada
3Clinical Epidemiology Program, Ottawa Hospital Research Institute and the Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada
4Departments of Medicine and Community Health Sciences, University of Calgary, Calgary, Alberta, Canada
5Department of Medicine, University of Alberta, Edmonton, Alberta, Canada
6Department of Geriatric Medicine, University of Toronto, Toronto, Ontario, Canada

Acknowledgements The authors thank the British Columbia Ministry of Health for requesting this systematic review and their useful feedback on the review conception. They thank Laure Perrier for conducting the literature searches, Becky Skidmore for peer reviewing the MEDLINE search strategy and Alissa Epworth for conducting the forward citation scanning and the PubMed related article searches. Finally, they thank Dr Maggie Chen and Dr Joseph Beyene for providing feedback on our original proposal, Judy Tran for locating full-text articles, Sophie Tsouros and Amir Zargarzadeh for screening some titles and abstracts, as well as Wing Hui for scanning the reference lists of some included studies, helping to generate the data tables and formatting the manuscript.

Contributors ACT conceived the study, designed the study, helped obtain funding for the study, screened literature for inclusion, abstracted data from included studies and wrote the manuscript. JA coordinated the study, screened the literature search results and abstracted data. JSH analysed the results and helped draft sections of the paper. PAK, MZ, HA and EB helped screen the literature and/or abstracted data. CS provided methods support. CHY, BH, BRH, DM and SRM helped obtain funding for the study and helped conceive the study. SES conceived the study, designed the study, obtained the funding and helped write the draft paper. All authors interpreted the results, read, edited and approved the final paper.

Funding This systematic review was funded by the Canadian Institutes for Health Research/Drug Safety and Effectiveness Network (CIHR/DSEN) (grant number 286941 DC01909P). ACT and BH are funded by CIHR/DSEN New Investigator Awards in Knowledge Synthesis. DM is funded by a University of Ottawa Research Chair. BRH is funded by a Roy and Vi Baay Chair in Kidney Research. SRM is the Endowed Chair in Patient Health Management (supported by the Faculties of Medicine and Dentistry and Pharmacy and Pharmaceutical Sciences) and holds a Health Scholar salary award (supported by Alberta Heritage Foundation for Medical Research and Alberta Innovates—Health Solutions). SES is funded by a Tier 1 Canada Research Chair in Knowledge Translation.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement The full data set is available, including data from the five included unpublished studies, on request from the corresponding author (SES, sharon.straus@utoronto.ca).

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

REFERENCES
32. Violante R, Oliveira JH, Yoon KH, et al. Randomized non-inferiority study comparing the addition of exenatide twice daily to sitagliptin or switching from sitagliptin to exenatide twice daily in patients with


