Dipeptidyl peptidase-4 inhibitors and risk of heart failure in type 2 diabetes: systematic review and meta-analysis of randomised and observational studies

Ling Li,1 Sheyu Li,2 Ke Deng,3 Jiali Liu,1 Per Olav Vandvik,4, 5 Pujing Zhao,1 Longhao Zhang,1 Jiantong Shen,1 Malgorzata M Bala,6 Zahra N Sohani,7, 8 Evelyn Wong,9 Jason W Busse,7 10, 11 Shanil Ebrahim,7 10, 12, 13 German Malaga,14 Lorena P Rios,15 Yingqiang Wang,16 Qunfei Chen,17 Gordon H Guyatt,2, 18 Xin Sun1

ABSTRACT

OBJECTIVES

To examine the association between dipeptidyl peptidase-4 (DPP-4) inhibitors and the risk of heart failure or hospital admission for heart failure in patients with type 2 diabetes.

DESIGN

Systematic review and meta-analysis of randomised and observational studies.

DATA SOURCES

Medline, Embase, Cochrane Central Register of Controlled Trials, and ClinicalTrials.gov searched up to 25 June 2015, and communication with experts.

ELIGIBILITY CRITERIA

Randomised controlled trials, non-randomised controlled trials, cohort studies, and case-control studies that compared DPP-4 inhibitors against placebo, lifestyle modification, or active antidiabetic drugs in adults with type 2 diabetes, and explicitly reported the outcome of heart failure or hospital admission for heart failure.

DATA COLLECTION AND ANALYSIS

Teams of paired reviewers independently screened for eligible studies, assessed risk of bias, and extracted data using standardised, pilot tested forms. Data from trials and observational studies were pooled separately; quality of evidence was assessed by the GRADE approach.

RESULTS

Eligible studies included 43 trials (n=68,775) and 12 observational studies (nine cohort studies, three nested case-control studies; n=1777,358). Pooling of 38 trials reporting heart failure provided low quality evidence for a possible similar risk of heart failure between DPP-4 inhibitor use versus control (42/15,701 vs 33/12,591; odds ratio 0.97 (95% confidence interval 0.61 to 1.56); risk difference 2 fewer (19 fewer to 28 more) events per 1000 patients with type 2 diabetes over five years). The observational studies provided effect estimates generally consistent with trial findings, but with very low quality evidence. Pooling of the five trials reporting admission for heart failure provided moderate quality evidence for an increased risk in patients treated with DPP-4 inhibitors versus control (622/18,554 vs 552/18,474; 1.13 (1.00 to 1.26); 8 more (0 more to 16 more)). The pooling of adjusted estimates from observational studies similarly suggested (with very low quality evidence) a possible increased risk of admission for heart failure (adjusted odds ratio 1.41, 95% confidence interval 0.95 to 2.09) in patients treated with DPP-4 inhibitors (exclusively sitagliptin) versus no use.

CONCLUSIONS

The relative effect of DPP-4 inhibitors on the risk of heart failure in patients with type 2 diabetes is uncertain, given the relatively short follow-up and low quality of evidence. Both randomised controlled trials and observational studies, however, suggest that these drugs may increase the risk of hospital admission for heart failure in those patients with existing cardiovascular diseases or multiple risk factors for vascular diseases, compared with no use.

Introduction

Of over 380 million people with diabetes worldwide, most (85–95%) have type 2 diabetes.1 Dipeptidyl peptidase-4 (DPP-4) inhibitors are a relatively new class of incretin based agents for treating type 2 diabetes. Evidence from randomised controlled trials has established that DPP-4 inhibitors reduce levels of glycated haemoglobin (HbA1c),2 3 do not affect body weight,2 4 pose a low risk of hypoglycaemia,2 4 and do not increase the risk of cardiovascular events.5 6 The American Diabetes Association and European Association for the Study of Diabetes have recommended this drug class as second line agents for type 2 diabetes management.8 A recent major trial9 (SAVOR-TIMI 53) reported an increased risk of admission to hospital for heart failure (hazard ratio 1.27, 95% confidence interval 1.07 to 1.51) with the DPP-4 inhibitor saxagliptin. Although unexpected, the finding raised concern among professionals and health authorities. In 2014, the US Food and Drug...
Administration (FDA) requested the clinical trial data from the manufacturer to investigate the potential association between use of saxagliptin and heart failure. The FDA then recommended that “Patients should not stop taking saxagliptin and should speak with their health care professionals about any questions or concerns. Health care professionals should continue to follow the prescribing recommendations in the drug labels.”

Subsequently, the EXAMINE trial testing alogliptin, and the TECOS trial testing sitagliptin, reported no significant effect on hospital admission for heart failure. Evidence from observational studies has been inconsistent,13-17 and the effect of DPP-4 inhibitors on heart failure remains controversial.

A systematic review of trials and observational studies offers an opportunity to consider the total body of evidence and potentially resolve the concern. We therefore undertook a systematic review to assess the extent to which DPP-4 inhibitors affect the risk of heart failure or hospital admission for heart failure in patients with type 2 diabetes.

Methods
We followed the standards set by the meta-analysis of observational studies in epidemiology (MOOSE)18 and preferred reporting items for systematic reviews and meta-analyses (PRISMA)19 for the reporting of our study.

Eligibility criteria
We included randomised controlled trials, non-randomised controlled trials, cohort studies, and case-control studies that compared DPP-4 inhibitors against placebo, lifestyle modification, or active antidiabetic drugs in adults with type 2 diabetes. We required follow-up for at least 12 weeks (not applicable to case-control studies), and explicit reporting of the outcome of heart failure or hospital admission for heart failure (either as raw data or adjusted effect estimates with 95% confidence intervals). We classified study designs according to recommendations by the Cochrane Non-Randomised Studies Methods Group. Trials, particularly phase III studies, reported heart failure either as a normal adverse event or a serious adverse event. For serious adverse events, admission for heart failure may have been included. We defined heart failure reported in such trials as an unspecified outcome.

Literature search
We searched Medline, Embase, and the Cochrane Central Register of Controlled Trials (CENTRAL) from inception to 25 June 2015. We combined both MeSH and free text terms for identifying relevant articles. An information expert (DP) developed our search strategies (web appendix 1).

We also searched ClinicalTrials.gov to identify additional eligible studies. Section 801 of the US Food and Drug Administration Amendments Act (FDAAA 801) requires responsible parties to submit summary results of clinical trials, including serious adverse events and adverse events with frequency over 5%, to this trial registry.20,21 In doing so, important information regarding heart failure can be collected. We used generic names of each drug to identify relevant studies, and limited our search to studies labelled as “completed” or “terminated,” in which summary results were available.

We also contacted content experts and industry representatives, and searched for conference abstracts on the American Diabetes Association and European Association for the Study of Diabetes, for additional information.

Study process
Teams of two paired reviewers, trained in health research methods, independently screened titles, abstracts, and full texts for eligibility; assessed risk of bias; and collected data from each eligible study, using standardised, pilot tested forms, together with detailed instructions. Reviewers resolved disagreement through discussion or, if required, by adjudication by a third reviewer (XS).

Risk of bias assessment
We used the Cochrane Collaboration’s tool22 to assess the risk of bias of randomised controlled trials. The items included random sequence generation, allocation concealment, blinding of participants, caregivers, and assessors of outcomes (that is, heart failure or hospital admission for heart failure), and adjudication of the outcomes. By assessing the risk of bias associated with blinding of patients, caregivers, and outcome assessors, we modified the instrument by removing the “unclear” option, an approach that we have previously shown to be reliable and valid.23

We used the Newcastle-Ottawa quality assessment scale25 to assess the risk of bias of cohort studies and case-control studies. We removed the items “representativeness of the exposed cohort” and “was follow-up long enough for outcomes to occur” for cohort studies and the item “representativeness of the cases” for case-control studies because these items relate to applicability of results. We, however, added two items: the ascertainment of type 2 diabetes and the ascertainment of potential confounding factors for these both types of studies, because misclassification could result from suboptimal approaches to these issues. We planned to assess publication bias but were unable to do so owing to very low event rates.

Data collection
We collected the following information from each eligible randomised controlled trial:

- General study characteristics: author name, year of publication, total number of patients randomised, number of treatment groups, length of follow-up, study phase, funding source, trial registry number, countries involved, and number of study sites
- Patient characteristics: sex, age, diabetes duration, body mass index, baseline HbA1c level, and fasting plasma glucose values
- Interventions: medications common to all groups (baseline treatment), details of DPP-4 inhibitors treatment and control group (eg, drug generic name, and duration of treatment)
• Outcomes: the definition of heart failure, number of events, and patients included for analyses in each group, as well as adjusted data if available.

For each trial, if the initial treatment assignment was switched (eg, patients in placebo group started receiving DPP-4 inhibitors after 24 weeks), we collected the outcome data before that point. If a trial had multiple reports, we collated all data into one study. If a trial had both reports from ClinicalTrials.gov and journal publications, we carefully checked data from these two sources for consistency. If outcome data were reported at multiple follow-up points, we used data from the longest follow-up.

For observational studies, we collected data similar to randomised controlled trials (eg, total number of patients, sex, age, diabetes duration, body mass index, baseline HbA1c). We documented, for each observational study, the definition of outcomes and sources of data for the outcomes. In addition, we documented information on:

• Study design (eg, retrospective or prospective cohort study)
• Data source (eg, claims data, electronic medical records)
• Methods used to ascertain type 2 diabetes status (eg, International Classification of Diseases (ICD) code)
• Exposures (eg, DPP-4 inhibitors, and other exposure variables)
• Methods used to control confounding (eg, logistic or cox regression, and control variables).

We collected adjusted estimates and their associated 95% confidence intervals, as well as adjustment factors, in addition to raw event data and exposure time.

Data analysis
We conducted separate analyses for randomised controlled trials and observational studies. We also separately analysed the data on heart failure and hospital admission for heart failure, because those two outcomes, although sharing the same clinical and pathophysiological features, represent differential seriousness of the effect of DPP-4 inhibitors treatment on patients and society. Heart failure could be subclinical and might not be diagnosed; admission for heart failure is, however, always a clinical event and a condition important to patients and clinicians. We considered admission for heart failure as the more important outcome for patients.

For the analysis of trials, we pooled outcome data using Peto’s methods because of very low event rates, and reported pooled Peto odds ratios and associated 95% confidence intervals. We examined heterogeneity among studies with the Cochrane $\chi^2$ test and the $I^2$ statistic. We explored sources of heterogeneity with four prespecified subgroup hypotheses:

• Type of control (placebo vs active treatment; larger effect in trials with placebo control)
• Length of follow-up ($\leq 52$ vs $>52$ weeks; larger effect in those with longer follow-up)
• Mode of treatment (monotherapy vs add-on or combination therapy; larger effect in those with add on or combination therapy)
• Individual DPP-4 inhibitors (different DPP-4 inhibitors vs control).

We carried out sensitivity analyses by using alternative effect measures (odds ratios vs risk ratios), pooling methods (Peto vs Mantel-Hanszel method), and statistical models regarding heterogeneity (random vs fixed effects).

In the analysis of observational studies, we qualitatively summarised the data for heart failure, because of the substantial variations in the comparison (that is, type of control) and patient populations in those studies. We pooled adjusted estimates of hospital admission for heart failure from cohort and nested case-control studies using a random effects model. Although the effect measures differ for those two designs (hazard ratios for cohort studies and odds ratios for nested case-control studies), they are relative measures and the effect estimates are close when the event rate is low (<5%).

Quality of evidence
We used the grading of recommendations assessment, development, and evaluation (GRADE) methodology to rate quality of the evidence for heart failure and hospital admission for heart failure as high, moderate, low, or very low. Randomised controlled trials begin as high quality evidence, but can be rated down because of risk of bias, imprecision, inconsistency, indirectness, and publication bias. Observational studies begin as low quality evidence, but can be rated up for a large magnitude of effect, a dose-response gradient, or presence of plausible confounders or other biases that increase confidence in the estimated effect.

Patient involvement
No patients were involved in setting the research question or the outcome measures, nor were they involved in developing plans for design or implementation of the study. No patients were asked to advise on interpretation or writing up of results. There are no plans to disseminate the results of the research to study participants or the relevant patient community.

Results
Of 11,440 potentially relevant reports identified, after title and abstract screening, 820 reports proved potentially eligible. On full text screening, 55 studies proved eligible, including 43 randomised controlled trials, representing 68,775 patients, reported in 77 reports (45 from journal reports, 31 from the ClinicalTrials.gov website, and one conference abstract) and 12 observational studies. Involving 1,777,358 patients, reported in nine cohort studies and three nested case-control studies (nine from journal reports, one from a trial registry, and two conference abstracts; fig 1). Two cohort studies analysed patient data from the same claims database, one presenting a subpopulation of the other; we included only the larger cohort study.
Evidence from randomised controlled trials

Trials reporting heart failure

Of the 43 randomised controlled trials, 38 reported heart failure, of which 33 (87%) were international studies, and 35 (92%) were clearly labelled as phase III trials. These 38 trials enrolled 109,278 patients (total n=31,680; mean age range 49.7-72.6 years, mean body mass index 24.0-32.8, mean baseline HbA1c 7.1-9.9%, mean fasting plasma glucose 7.7-11.1 mmol/L, and mean duration of diabetes 1.7-17.5 years; table 1). Nine trials used DPP-4 inhibitors as monotherapy, 27 as add-on or combination therapy, and two as both monotherapy and combination therapy. Length of follow-up ranged from 12 to 206 weeks (median 52; table 2).

All 38 trials were industry funded. Most (n=24) were identified from ClinicalTrials.gov, of which four (9-93%) have not been published in a peer reviewed journal. Because of the limited information in the trial registry, we were unable to adequately assess the risk of bias for these four trials. On the basis of the information we collected, 16 (42%) trials adequately generated their randomisation sequence; 11 (29%) adequately concealed allocation; all trials blinded patients, caregivers, and outcome assessors; eight (21%) adjudicated heart failure events; and four (11%) used blinded assessors to adjudicate heart failure (web appendix 2). The treatment groups of each included trial were generally balanced with respect to demographic and clinical characteristics.

Effects on heart failure

The 38 trials reported 75 heart failure events occurring in 28,292 patients who were treated with at least one drug (raw event rate 0.27%). The definition of heart failure was available in only one trial; 37 33 (87%) trials reported heart failure as serious adverse events. The pooling of data from these trials showed no significant difference in the risk of heart failure between DPP-4 inhibitors treatment and control. Event rates were 0.27% for DPP-4 inhibitors versus 0.26% for controls (odds ratio 0.97 (95% confidence interval 0.61 to 1.56), I²=60%; risk difference 2 fewer (19 fewer to 28 more) events per 1000 patients with type 2 diabetes over five years; fig 2 and table 3). We rated the quality of evidence as low because of risk of bias and imprecision (table 3).

The subgroup analysis by type of control (placebo vs active drugs) showed no difference in treatment effects (interaction P=0.57; comparison with placebo, odds ratio 1.17 (95% confidence interval 0.58 to 2.33); comparison with active drugs, 0.89 (0.47 to 1.66); fig A in web appendix 3). The subgroup analyses of the other three prespecified hypotheses showed no difference in treatment effects (figs B-D in appendix 3). Sensitivity analysis using alternative effect measures, statistical methods, and analysis models did not show important changes in pooled effects (figs E-G in web appendix 3).

Trials reporting hospital admission for heart failure

We included three large trials9 11 12 (SAVOR-TIMI 53, EXAMINE, and TECOS) and two small trials104 105 reporting hospital admission for heart failure; all were designed to assess the cardiovascular safety of DPP-4 inhibitors compared with placebo (table 1). The SAVOR-TIMI 53 trial investigated saxagliptin in patients with diabetes who had a renal impairment and cardiovascular disease or multiple risk factors for vascular disease. The EXAMINE trial recruited patients receiving alogliptin with type 2 diabetes and a recent acute coronary syndrome. The TECOS trial examined sitagliptin in patients with type 2 diabetes and cardiovascular disease. In addition, one small trial106 assessed vildagliptin in patients with type 2 diabetes as well as heart failure and a left ventricular ejection fraction less than 40%; the other small trial107 assessed linagliptin in patients with type 2 diabtes with moderate to severe renal impairment.

All three large trials were international studies. The median length of follow-up ranged from 76 to 156 weeks (table 1). Those trials enrolled 5380-16,492 patients (total n=36,607; mean age range 60.9-65.5 years, mean body mass index 29.5-31.1, and duration of diabetes 9.2-11.6 years). The two small trials followed up patients for 52 weeks; mean age ranged from 63 to 66.6 years and mean HbA1c levels ranged from 7.8% to 8.1%.

All trials, but one106 (which had unclear details because it was presented as an abstract), adequately generated their randomisation sequence and adequately concealed allocation; all trials blinded patients, caregivers, outcome assessors, and centrally adjudicated hospital admission for heart failure outcome through a clinical events classification committee who were blinded to treatment allocation. All trials were funded by industry (web appendix 2).

Effects on hospital admission for heart failure

All five trials9 11 12 104 105 reported unadjusted rates of hospital admission for heart failure. Overall, 1174 events of heart failure occurred in 37,028 patients
Table 1: Characteristics of included randomised controlled trials

<table>
<thead>
<tr>
<th>Trials reporting heart failure</th>
<th>Author (year)</th>
<th>International study</th>
<th>No of countries involved</th>
<th>No of study sites</th>
<th>Study phase</th>
<th>Total No of patients randomised</th>
<th>Length of follow-up (weeks)</th>
<th>Male patients (No, %)</th>
<th>Mean age (years)</th>
<th>Mean body mass index</th>
<th>Mean HbA1c (%)</th>
<th>Mean FPG (mmol/L)</th>
<th>Mean diabetes duration (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arjona Ferreira (2013)a 13-17</td>
<td>Yes</td>
<td>NR</td>
<td>NR</td>
<td>III</td>
<td>426</td>
<td>54</td>
<td>158 (57)</td>
<td>64.5</td>
<td>26.8</td>
<td>7.8</td>
<td>8.1</td>
<td>10.4</td>
<td></td>
</tr>
<tr>
<td>Bosi (2010)b,c</td>
<td>Yes</td>
<td>NR</td>
<td>NR</td>
<td>III</td>
<td>801</td>
<td>52</td>
<td>414 (11.6)</td>
<td>55.1</td>
<td>31.5</td>
<td>8.2</td>
<td>9.0</td>
<td>7.2</td>
<td></td>
</tr>
<tr>
<td>Ferrannini (2009)c,k,l</td>
<td>Yes</td>
<td>NR</td>
<td>NR</td>
<td>III</td>
<td>833</td>
<td>52</td>
<td>158 (57)</td>
<td>64.5</td>
<td>26.8</td>
<td>7.8</td>
<td>8.1</td>
<td>10.4</td>
<td></td>
</tr>
<tr>
<td>Garber (2007)d</td>
<td>Yes</td>
<td>NR</td>
<td>NR</td>
<td>III</td>
<td>615</td>
<td>54</td>
<td>972 (56.5)</td>
<td>30.9</td>
<td>8.0</td>
<td>10.8</td>
<td>5.4</td>
<td>7.9</td>
<td></td>
</tr>
<tr>
<td>Iwamoto (2010)i,a,l,m,n</td>
<td>Yes</td>
<td>NR</td>
<td>NR</td>
<td>III</td>
<td>722</td>
<td>12</td>
<td>224 (61.7)</td>
<td>59.8</td>
<td>34.5</td>
<td>7.6</td>
<td>8.3</td>
<td>5.4</td>
<td></td>
</tr>
<tr>
<td>Scirica (2013)m</td>
<td>Yes</td>
<td>NR</td>
<td>NR</td>
<td>III</td>
<td>132</td>
<td>12</td>
<td>254 (75.6)</td>
<td>82.4</td>
<td>12.7</td>
<td>9.3</td>
<td>8.2</td>
<td>5.4</td>
<td></td>
</tr>
<tr>
<td>Pratley (2014)c</td>
<td>Yes</td>
<td>NR</td>
<td>NR</td>
<td>III</td>
<td>126</td>
<td>12</td>
<td>254 (75.6)</td>
<td>82.4</td>
<td>12.7</td>
<td>9.3</td>
<td>8.2</td>
<td>5.4</td>
<td></td>
</tr>
<tr>
<td>Pratley (2014)d</td>
<td>Yes</td>
<td>NR</td>
<td>NR</td>
<td>III</td>
<td>126</td>
<td>12</td>
<td>254 (75.6)</td>
<td>82.4</td>
<td>12.7</td>
<td>9.3</td>
<td>8.2</td>
<td>5.4</td>
<td></td>
</tr>
<tr>
<td>Green (2015)l,e</td>
<td>Yes</td>
<td>NR</td>
<td>NR</td>
<td>III</td>
<td>126</td>
<td>12</td>
<td>254 (75.6)</td>
<td>82.4</td>
<td>12.7</td>
<td>9.3</td>
<td>8.2</td>
<td>5.4</td>
<td></td>
</tr>
<tr>
<td>Krum (2014)m,a</td>
<td>Yes</td>
<td>NR</td>
<td>NR</td>
<td>III</td>
<td>126</td>
<td>12</td>
<td>254 (75.6)</td>
<td>82.4</td>
<td>12.7</td>
<td>9.3</td>
<td>8.2</td>
<td>5.4</td>
<td></td>
</tr>
<tr>
<td>Look (2015)c,d</td>
<td>Yes</td>
<td>NR</td>
<td>NR</td>
<td>III</td>
<td>126</td>
<td>12</td>
<td>254 (75.6)</td>
<td>82.4</td>
<td>12.7</td>
<td>9.3</td>
<td>8.2</td>
<td>5.4</td>
<td></td>
</tr>
<tr>
<td>Scirica (2013)m</td>
<td>Yes</td>
<td>NR</td>
<td>NR</td>
<td>III</td>
<td>126</td>
<td>12</td>
<td>254 (75.6)</td>
<td>82.4</td>
<td>12.7</td>
<td>9.3</td>
<td>8.2</td>
<td>5.4</td>
<td></td>
</tr>
<tr>
<td>Zamnad (2015)l</td>
<td>Yes</td>
<td>NR</td>
<td>NR</td>
<td>III</td>
<td>126</td>
<td>12</td>
<td>254 (75.6)</td>
<td>82.4</td>
<td>12.7</td>
<td>9.3</td>
<td>8.2</td>
<td>5.4</td>
<td></td>
</tr>
</tbody>
</table>

FPG=fasting plasma glucose; NR=not reported.
*Median diabetes duration (years).
**No (%) of patients with no more than five years’ duration.
#Median follow-up time (weeks).

(raw event rate 3.4% for DPP-4 inhibitors v 3.0% for controls; table 3). Pooling across trials showed a borderline increase in the risk of hospital admission for heart failure in patients with type 2 diabetes using DPP-4 inhibitors versus control (odds ratio 1.13 (95% confidence interval 1.00 to 1.26), P=0.01; risk difference 8 more (0 more to 16 more) per 1000 patients with type 2 diabetes over five years; fig 1 and table 3). We rated the quality of evidence as moderate due to imprecision (table 3). Sensitivity analysis by use of alternative effect measures, statistical methods, and analysis models did not show important changes in the pooled effects (figs H-J in web appendix 3).

Evidence from observational studies
Of 12 observational studies, four109-112 reported heart failure, and eight113-115 reported hospital admission for heart failure; nine105-113,115 were cohort studies and the other three116,117 were nested case-control studies (fig 1).
### Observational studies reporting heart failure

Of the four studies reporting heart failure, two prospective cohort studies compared DPP-4 inhibitors versus sulfonylureas and sitagliptin versus sulfonylureas. One retrospective cohort study assessed DPP-4 inhibitors versus sulfonylureas and reported the findings from the subgroup of DPP-4 inhibitors. Finally, one nested case-control study using claims data investigated use of sitagliptin versus no use in patients admitted to hospital for acute coronary syndrome.

### Table 2 | Interventions tested and event rates in randomised controlled trials

<table>
<thead>
<tr>
<th>Author (year)</th>
<th>Drug treatments used across groups</th>
<th>DPP-4 inhibitors</th>
<th>Control</th>
<th>Events/analysed patients (No)</th>
<th>Duration of treatment (weeks)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arjona Ferreira (2013)a</td>
<td>None</td>
<td>Sitagliptin 0/210</td>
<td>Glipizide 4/212</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>Arjona Ferreira (2013)b</td>
<td>None</td>
<td>Sitagliptin 2/64</td>
<td>Glipizide 2/65</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>Bosi (2013)</td>
<td>Metformin, and pioglitazone 30 mg</td>
<td>Alogliptin 2/404</td>
<td>Add-on pioglitazone 15 mg 1/399</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td>Ferrannini (2009)</td>
<td>Metformin</td>
<td>Vildagliptin 2/1389</td>
<td>Glimepiride 2/1383</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td>Fonseca (2013)</td>
<td>Metformin and pioglitazone</td>
<td>Sitagliptin 0/157</td>
<td>Placebo 0/156</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>Henry (2014)</td>
<td>Sitagliptin</td>
<td>Vildagliptin 2/691</td>
<td>No additional drugs 0/693</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>Iwamoto (2010)</td>
<td>Sitagliptin</td>
<td>Alogliptin 1/290</td>
<td>Placebo 0/73</td>
<td>12</td>
<td></td>
</tr>
</tbody>
</table>

### Trials reporting hospital admission for heart failure

<table>
<thead>
<tr>
<th>Author (year)</th>
<th>Drug treatments used across groups</th>
<th>DPP-4 inhibitors</th>
<th>Control</th>
<th>Events/analysed patients (No)</th>
<th>Duration of treatment (weeks)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arjona Ferreira (2013)a</td>
<td>None</td>
<td>Sitagliptin 0/210</td>
<td>Glipizide 4/212</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>Arjona Ferreira (2013)b</td>
<td>None</td>
<td>Sitagliptin 2/64</td>
<td>Glipizide 2/65</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>Bosi (2013)</td>
<td>Metformin, and pioglitazone 30 mg</td>
<td>Alogliptin 2/404</td>
<td>Add-on pioglitazone 15 mg 1/399</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td>Ferrannini (2009)</td>
<td>Metformin</td>
<td>Vildagliptin 2/1389</td>
<td>Glimepiride 2/1383</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td>Fonseca (2013)</td>
<td>Metformin and pioglitazone</td>
<td>Sitagliptin 0/157</td>
<td>Placebo 0/156</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>Henry (2014)</td>
<td>Sitagliptin</td>
<td>Vildagliptin 2/691</td>
<td>No additional drugs 0/693</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>Iwamoto (2010)</td>
<td>Sitagliptin</td>
<td>Alogliptin 1/290</td>
<td>Placebo 0/73</td>
<td>12</td>
<td></td>
</tr>
</tbody>
</table>

### Observational studies reporting heart failure

Of the four studies reporting heart failure, two prospective cohort studies compared DPP-4 inhibitors versus sulfonylureas and sitagliptin versus sulfonylureas. One retrospective cohort study assessed DPP-4 inhibitors versus sulfonylureas and reported the findings from the subgroup of DPP-4 inhibitors. Finally, one nested case-control study using claims data investigated use of sitagliptin versus no use in patients admitted to hospital for acute coronary syndrome.
cians in one prospective cohort study. The diagnostic criteria for heart failure ranged from one to four years. Enrolled patients had a risk of heart failure in patients with type 2 diabetes who received DPP-4 inhibitors versus control from randomised controlled trials.

Four studies used registry data, electronic health or medical records, or claims data for their analyses. Patients with type 2 diabetes were ascertained by physicians in one prospective cohort study or by ICD-9 Clinical Modification (CM) codes in one nested case-control study; the other two cohort studies did not explicitly report the ascertainment of type 2 diabetes. None of these studies mentioned the ascertainment of exposure to DPP-4 inhibitors agents and other confounding variables; the accuracy of ascertaining exposure and confounding factors was unclear. Of these three cohort studies, only one demonstrated that the outcome of interest was not present at start of study, and mentioned the method used to assess the outcome of interest. Of these four studies, two controlled for the effect of confounding factors (web appendices 4 and 5).

**Effects on heart failure**
All three cohort studies reported unadjusted rates of heart failure, involving 541 events among 16 408 patients (raw event rate 3.3%). Because of the

![Fig 2 | Risk of heart failure in patients with type 2 diabetes who received DPP-4 inhibitors versus control from randomised controlled trials](image-url)
heterogeneous and indirect nature of the identified evidence—with substantial variations in comparisons and types of patients—we did not pool data across studies. The outcome information is presented in table 5.

One retrospective cohort study\(^\text{111}\) and one nested case-control study\(^\text{112}\) reported adjusted data. The retrospective cohort study, including 13 185 patients with a median follow-up of four years, suggested that, compared with sulfonylureas, DPP-4 inhibitors was statistically associated with an increased risk of congestive heart failure (adjusted hazard ratio 1.10, 95% confidence interval 1.04 to 1.17). The nested case-control study, selecting 457 heart failure cases and 4 570 controls, showed no statistical difference in the risk of heart failure between use and no use of sitagliptin the 90 days before acute coronary syndrome (adjusted odds ratio 0.75, 95% confidence interval 0.38 to 1.46, table 5). Using GRADE, we rated the quality of evidence in the identified studies as very low, owing to risk of bias, indirectness, and imprecision in addition to the inherent risk for confounding given the observational design.

**Observational studies reporting hospital admission for heart failure**

Of the eight studies reporting hospital admission for heart failure, six retrospective cohort studies,\(^\text{13-15,113-115}\) using registry data, claims data, or electronic medical records, assessed DPP-4 inhibitors versus active drugs (eg, sulfonylureas, pioglitazone), and the use of sitagliptin versus no use of sitagliptin. The other two nested case-control studies\(^\text{16,17}\) assessed use of sitagliptin versus no use of sitagliptin, and incretin based drugs (including the DPP-4 inhibitors subgroup) versus other oral anti-diabetic drugs (tables 4 and 5).

The sample sizes of these eight studies ranged from 4 224 to 9 351, and the mean or median length of follow-up ranged from 0.5 to 2.6 years. Enrolled patients had a mean age ranging from 54.6 to 67 years, mean baseline HbA1c level ranging from 7.5% to 8.0%, and a mean duration of diabetes ranging from 2.3 to 8.6 years.

The eight studies used registry data, claims data, or electronic medical records for analyses. Only two studies\(^\text{15,111}\) explicitly reported use of ICD codes to ascertain patients with diabetes; one study\(^\text{15}\) ascertained exposure to DPP-4 inhibitors by using anatomical therapeutic chemical classes; three studies\(^\text{13,15,115}\) explicitly stated use of ICD codes to ascertain other confounding variables. Four studies\(^\text{13-15,113,115}\) used ICD-9 or ICD-10 codes to assess outcomes. Three cohort studies\(^\text{13,15,115}\) clarified that the outcome of interest was not present at enrolment. All eight studies controlled for potential confounding factors, but failed to specify the extent to which the data were complete in the database (web appendices 4 and 5).

**Effects on hospital admission for heart failure**

All but one retrospective cohort study\(^\text{115}\) reported unadjusted rates of hospital admission for heart failure. The five cohort studies\(^\text{13-15,113,115}\) included 3 500 events among 1 630 884 patients (raw event rate 0.2%;
1466 events (0.2%) in 912,309 patients from the DPP-4 inhibitors group, and 2034 events (0.3%) in 718,575 patients from the control group). The two nested case-control studies involved 1942 cases among 27,806 patients. Because of the variety of confounding factors investigated in the studies, we did not pool the unadjusted data.

All eight studies reported adjusted estimates of hospital admission for heart failure. Of these, six studies—five cohort studies and one nested case-control study—compared DPP-4 inhibitors with active drugs (sulfonylureas, pioglitazone, other oral antidiabetic drugs). Pooling of adjusted estimates from these six studies showed that DPP-4 inhibitors were associated with reduced risk of hospital admission for heart failure (adjusted odds ratio 0.85, 95% confidence interval 0.74 to 0.97; I²=31%). However, pooling of the cohort study (16,576 patients and 614 events), and the nested case-control study (82% cases and 8238 controls) suggested a non-significant trend for increased risk of admission for heart failure compared with no use of sitagliptin (adjusted odds ratio 1.41, 0.95 to 2.09; I²=65%). There was significant subgroup effect by type of control (interaction P=0.02, fig 4). Using GRADE, we rated the quality of evidence as very low, due to risk of bias, heterogeneity, and imprecision in addition to the inherent risk for confounding given the observational design.

Table 6 summarises the evidence regarding the effects of DPP-4 inhibitors on heart failure or hospital admission for heart failure.

**Discussion**

**Main findings**

The only evidence of moderate quality from our results is from randomised controlled trials that examined the effect of DPP-4 inhibitors on hospital admission for heart failure. These studies suggested a small increase, in both relative and absolute terms, in heart failure admissions in patients using DPP-4 inhibitors than those not. The results, however, are of borderline significance. Evidence from observational studies is of very low quality, and thus has little bearing on any inferences about DPP-4 inhibitor effects on heart failure admission.

With respect to the incidence of heart failure, trial evidence leaves uncertainty regarding the relative effect of DPP-4 inhibitors. Because the follow-up was relatively short and the baseline risk of patients was very low in those trials, the incidence of heart failure was very low (well under 1% per year), and with the small number of events, the confidence intervals around relative effects are wide. In addition, heart failure was unspecified in all but one of the phase III trials. Many (87%) reported heart failure as serious adverse events, in which admission for heart failure might have been included according to the definition of serious adverse events. The pooled estimate could thus represent a composite of heart failure with or without admission for heart failure. The observational studies again provide very low quality evidence and have little effect on inferences, although results are consistent. Overall, the current evidence provides no support for the hypothesis that DPP-4 inhibitors increase the incidence of heart failure.

**Strengths and limitations**

Our study has several strengths. Firstly, we used rigorous methods to systematically identify and include data from both randomised and non-randomised studies to examine the effect of DPP-4 inhibitors on risk of heart failure and hospital admission for heart failure. Secondly, in addition to published reports, we have identified additional data from ClinicalTrials.gov. Our study included four randomised controlled trials and three observational studies that were not published in journals. Thirdly, we instituted a rigorous approach to ensure the data were accurate. In particular, we carefully checked the data reported in ClinicalTrials.gov and journal publications for consistency. Fourthly, we addressed several prespecified subgroup analyses to explore sources of heterogeneity. Finally, we used GRADE to assess the quality of the body of evidence.

Our study also had some limitations. Firstly, for various reasons, some trials are likely not to report outcome data in their full publications. However, we have obtained additional data through the search of the ClinicalTrials.gov and conference abstracts, which minimised the risk of outcome reporting bias. Secondly, given the limitations of reported data, we were...
### Table 4 | Characteristics of included observational studies

<table>
<thead>
<tr>
<th>Study design reporting heart failure</th>
<th>Study design</th>
<th>Data source</th>
<th>Countries</th>
<th>Funding</th>
<th>Total No of patients</th>
<th>Follow-up (years)</th>
<th>Male patients (No (%))</th>
<th>Mean age (years)</th>
<th>Mean body mass index</th>
<th>Mean HbA1c (%)</th>
<th>Mean FPG (mmol/L)</th>
<th>Mean diabetes duration (years)</th>
<th>CVD at baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gitt (2013)109 Prospective cohort study</td>
<td>Registry data</td>
<td>Germany</td>
<td>Private for-profit funding</td>
<td>616</td>
<td>1</td>
<td>5097 (49.8)</td>
<td>65.8†</td>
<td>NR</td>
<td>7.3†</td>
<td>7.7†</td>
<td>4.9†</td>
<td>Included patients had CVD or had no CVD at baseline</td>
<td></td>
</tr>
<tr>
<td>NCT01357135 (2014)110 Prospective cohort study</td>
<td>Electronic medical records</td>
<td>France</td>
<td>Private for-profit funding</td>
<td>3453</td>
<td>3</td>
<td>2004 (58.0)</td>
<td>63.5</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>Kannan (2015)*,111 Retrospective cohort study</td>
<td>Electronic health records</td>
<td>USA</td>
<td>No funding</td>
<td>13185</td>
<td>4†</td>
<td>7827 (54.6)</td>
<td>60.6</td>
<td>32.6†</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>Included patients had no history of CVD or congestive heart failure at baseline</td>
</tr>
<tr>
<td>Eunch (2014)112 Nested case-control study</td>
<td>Claims data</td>
<td>USA</td>
<td>NR</td>
<td>5027</td>
<td>NA</td>
<td>3268 (65)</td>
<td>65</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>Included patients had no history of heart failure in the 3 years before admission to hospital for an acute coronary syndrome event</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study design reporting hospital admission for heart failure</th>
<th>Study design</th>
<th>Data source</th>
<th>Countries</th>
<th>Funding</th>
<th>Total No of patients</th>
<th>Follow-up (years)</th>
<th>Male patients (No (%))</th>
<th>Mean age (years)</th>
<th>Mean body mass index</th>
<th>Mean HbA1c (%)</th>
<th>Mean FPG (mmol/L)</th>
<th>Mean diabetes duration (years)</th>
<th>CVD at baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fadini (2015)13 Retrospective cohort study</td>
<td>Registry data</td>
<td>Italy</td>
<td>Public funding</td>
<td>127555</td>
<td>2.6</td>
<td>66201 (51.9)</td>
<td>67.0</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>Included patients had CVD or no CVD at baseline</td>
</tr>
<tr>
<td>Fu (2015)14 Retrospective cohort study</td>
<td>Claims data</td>
<td>USA</td>
<td>NR</td>
<td>218556</td>
<td>0.5</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>Included patients had CVD or no CVD at baseline</td>
</tr>
<tr>
<td>Seong (2015)115 Retrospective cohort study</td>
<td>Claims data</td>
<td>South Korea</td>
<td>No funding</td>
<td>349476</td>
<td>0.6</td>
<td>191167 (54.7)</td>
<td>58.3</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>Included patients had no history of CVD within 2.5 years before cohort entry</td>
</tr>
<tr>
<td>Suh (2015)116 Retrospective cohort study</td>
<td>Claims data</td>
<td>South Korea</td>
<td>NR</td>
<td>935519</td>
<td>0.9</td>
<td>518614 (55.4)</td>
<td>59.4</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>Velez (2015)*,117 Retrospective cohort study</td>
<td>Electronic medical records</td>
<td>USA</td>
<td>Public funding</td>
<td>4224</td>
<td>2.0†</td>
<td>2263 (53.6)</td>
<td>60.8</td>
<td>NR</td>
<td>8.0</td>
<td>NR</td>
<td>2.5</td>
<td>Included patients had CVD or no CVD at baseline</td>
<td></td>
</tr>
<tr>
<td>Wang (2014)118 Retrospective cohort study</td>
<td>Claims data</td>
<td>Taiwan</td>
<td>Public funding</td>
<td>46576</td>
<td>1.5†</td>
<td>8615 (52.0)</td>
<td>64.3</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>8.6</td>
<td>Included patients had CVD or no CVD at baseline</td>
<td></td>
</tr>
<tr>
<td>Wei (2014)119 Nested case-control study</td>
<td>Claims data</td>
<td>USA</td>
<td>NR</td>
<td>45434</td>
<td>NA</td>
<td>27013 (59.5)</td>
<td>54.6</td>
<td>NR</td>
<td>7.5</td>
<td>NR</td>
<td>NR</td>
<td>Included patients were recently diagnosed with heart failure</td>
<td></td>
</tr>
<tr>
<td>Yu (2015)*,120 Nested case-control study</td>
<td>Electronic medical records</td>
<td>UK</td>
<td>Public funding</td>
<td>57737</td>
<td>NA</td>
<td>32795 (56.8)</td>
<td>61.6</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>2.3</td>
<td>Included patients had CVD or no CVD at baseline</td>
<td></td>
</tr>
</tbody>
</table>

FPG=fasting plasma glucose; CVD=cardiovascular disease; NR=not reported; NA=not applicable.

*Three studies accessed incretin agents (both glucagon-like peptide-1 receptor agonists and DPP-4 inhibitors) and the risk of heart failure, so the data above were the characteristics of total patients included.

†Median value.
Table 5 | Exposures, outcomes, and results of included observational studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Exposure of interest</th>
<th>Control group</th>
<th>No of events or cases</th>
<th>Total no of analysed patients</th>
<th>Adjusted estimates (95% CI)</th>
<th>Adjusted covariate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eurich (2014)</td>
<td>Sitagliptin use</td>
<td>No use</td>
<td>457</td>
<td>5027</td>
<td>OR 0.75 (0.38 to 1.46)</td>
<td>Demographics, clinical and laboratory data, pharmacy claims, health care use and propensity scores (conditional probability of being treated with metformin or sulfonylurea or insulin or sitagliptin)</td>
</tr>
<tr>
<td>Kannan (2015)</td>
<td>DPP-4 inhibitors (combined with metformin)</td>
<td>Sulfonylureas (combined with metformin)</td>
<td>528*</td>
<td>13,185 (55.110 person years)*</td>
<td>HR 1.10 (1.04 to 1.17)</td>
<td>Age, sex, race, body mass index, number of encounters, median household income, smoking status, systolic and diastolic blood pressure, hypertension, dyslipidaemia, cerebral vascular event, presence of neuropathy, retinopathy, dementia, chronic obstructive pulmonary disease, cancer, atrial fibrillation, antihypertensive drugs, lipid lowering agents, antithrombotic agents, and propensity for being on metformin and sulfonylureas at baseline, lipid profile, estimated glomerular filtration rate</td>
</tr>
<tr>
<td>Gitt (2013)</td>
<td>DPP-4 inhibitors</td>
<td>Sulfonylureas</td>
<td>11</td>
<td>616</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>NCT01357135 (2014)</td>
<td>Sitagliptin (combined with metformin)</td>
<td>Sulfonylureas (combined with metformin)</td>
<td>2</td>
<td>2607</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Fadini (2015)</td>
<td>DPP-4 inhibitors</td>
<td>Sulfonylureas</td>
<td>1181</td>
<td>110,757</td>
<td>HR 0.78 (0.62 to 0.97)</td>
<td>Age, sex, use of certain medications (drugs for hypertension, dyslipidaemia, chronic obstructive pulmonary disease, non-steroidal anti-inflammatory drugs, and antithrombotic drugs), presence of previous hospital admissions, Charlson index level grouped into three categories, previous use of oral glucose lowering drugs, cotreatment with metformin, and adherence level categorised on the basis of the medication possession ratio (MPR (%), &lt;80% v≥80%)</td>
</tr>
<tr>
<td>Fu (2015)</td>
<td>DPP-4 inhibitors</td>
<td>Sulfonylureas</td>
<td>495</td>
<td>218,556</td>
<td>No CVD at baseline HR 0.59 (0.38 to 0.88), CVD at baseline: HR 0.78 (0.57 to 1.05)</td>
<td>Adjusted covariates of Cox proportional hazard models were not stated explicitly; each comparison consisted of patients matched 1:1 on a propensity score based on demographics, general clinical characteristics, and hospital admission for heart failure risk factors from one year before baseline, analyses were stratified by presence of CVD</td>
</tr>
<tr>
<td>Seong (2015)</td>
<td>DPP-4 inhibitors</td>
<td>Sulfonylureas and pioglitazone</td>
<td>212</td>
<td>349/46 (211,959 person years)</td>
<td>DPP-4 inhibitors vs sulfonylureas: adjusted HR 0.93 (0.62 to 1.41); DPP-4 inhibitors vs pioglitazone: 0.21 (0.15 to 0.28)</td>
<td>Adjusted factors included age, sex, duration of diabetes at baseline; comorbidities in year before the index date (microvascular complications of diabetes (retinopathy, neuropathy, or nephropathy), peripheral vascular disease, hypertension, and dyslipidaemia), and associated Charlson score, diabetes-related hospital admission and total number of hypoglycaemic drug classes used in year before the index date; and use of the following drug classes in year before the index date: hypoglycaemic, lipid lowering, antihypertensive, antithrombotic (drug names not listed here)</td>
</tr>
<tr>
<td>Suh (2015)</td>
<td>DPP-4 inhibitors</td>
<td>Pioglitazone</td>
<td>998</td>
<td>935,519</td>
<td>Sitagliptin vs pioglitazone: adjusted HR 0.97 (0.81 to 1.16), vildagliptin vs pioglitazone: 1.22 (0.99 to 1.50)</td>
<td>Age and sex</td>
</tr>
<tr>
<td>Velez (2015)</td>
<td>DPP-4 inhibitors</td>
<td>Control (no details)</td>
<td>127</td>
<td>3987</td>
<td>HR 0.58 (0.38, 0.88)</td>
<td>Propensity score, number of antidiabetic drugs, duration of diabetes, baseline beta blocker use, and use of angiotensin converting enzyme inhibitor or angiotensin receptor blocker</td>
</tr>
<tr>
<td>Wang (2014)</td>
<td>Sitagliptin use</td>
<td>No use</td>
<td>614</td>
<td>16,976</td>
<td>HR 1.21 (1.04 to 1.42)</td>
<td>Adjusted covariates of Cox proportional hazard models were not stated explicitly; potential confounding were mitigated by the propensity score matching approach, and covariates included age, sex, duration of diabetes, antidiabetic drugs used, comorbidities, and outpatient visit</td>
</tr>
<tr>
<td>Wei (2014)</td>
<td>Sitagliptin use</td>
<td>No use</td>
<td>824</td>
<td>9,062</td>
<td>OR 1.84 (1.16 to 2.92)</td>
<td>Demographics (age, sex, and socioeconomic status), most recent clinical laboratory data (HbA1c, low and high density lipoprotein cholesterol, triglycerides, estimated glomerular filtration rate, albuminuria, and haemoglobin A1c levels), history of CVD (ischaemic heart disease, myocardial infarction, dyslipidaemia, hypertension, arrhythmia, and valve disease), and prescription drug use (antiplatelet drugs, anticoagulants, statins, calcium channel blockers, β blockers, angiotensin converting enzyme inhibitors, renin inhibitors, diuretics, and nitrates)</td>
</tr>
<tr>
<td>Yu (2015)</td>
<td>DPP-4 inhibitors (sitagliptin, vildagliptin, and saxagliptin, alone or in combination with other antidiabetic drugs)</td>
<td>Other oral antidiabetic drugs</td>
<td>1118*</td>
<td>18 7%, 4*</td>
<td>OR 0.88 (0.63 to 1.22)</td>
<td>Sex, body mass index, excessive alcohol use, smoking status, HbA1c level, comorbidities (neuropathy, renal disease, retinopathy, atrial fibrillation, cancer (other than non-melanoma skin cancer), chronic obstructive pulmonary disease, coronary artery disease, dyslipidaemia, hypertension, previous myocardial infarction, peripheral arterial disease, previous coronary revascularisation, peripheral vascular disease, and previous stroke), number of prescriptions, number of physician visits, and use of the following drugs in the year before cohort entry: angiotensin converting enzyme inhibitors, angiotensin receptor blockers, beta blockers, calcium channel blockers, diuretics, fibrates, statins, aspirin, and other non-steroidal anti-inflammatory drugs</td>
</tr>
</tbody>
</table>

NR = not reported; HR = hazard ratio; OR = odds ratio; CVD = cardiovascular disease.

*These two studies accessed incretin drugs and the risk of heart failure, and data of events/cases and total number of analysed patients regarding glucagon-like peptide 1 receptor agonists and DPP-4 inhibitors were not reported separately, so the data above were of total study patients.
unable to confirm whether the increased risk of hospital admission for heart failure was a class effect or a specific effect of saxagliptin. Other limitations included those of the primary studies, such as the risk of bias of observational studies, the potentially variable specification of outcomes (heart failure and hospital admission for heart failure), and the likelihood of variable and incomplete ascertainment of heart failure in the clinical trials.

Comparison with other studies
Four previous meta-analyses have explored the effect of DPP-4 inhibitors on the risk of heart failure. Of those studies, one found that treatment with DPP-4 inhibitors for 29 weeks or longer was associated with an increased risk of new onset of heart failure (risk ratio 1.16, 95% confidence interval 1.01 to 1.33), but not with treatment for less than 29 weeks (0.67, 0.32 to 1.04). The second included 24 randomised controlled trials that

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>Log (odds ratio)</th>
<th>SE</th>
<th>Weight (%)</th>
<th>Odds ratio, IV Random (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DPP-4 inhibitors v active drugs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fadini 2015</td>
<td>-0.249</td>
<td>0.114</td>
<td>21.9</td>
<td>0.78 (0.62 to 0.98)</td>
</tr>
<tr>
<td>Fu 2015</td>
<td>-0.139</td>
<td>0.090</td>
<td>28.5</td>
<td>0.87 (0.73 to 1.04)</td>
</tr>
<tr>
<td>Seong 2015</td>
<td>-0.821</td>
<td>0.750</td>
<td>0.8</td>
<td>0.44 (0.10 to 1.91)</td>
</tr>
<tr>
<td>Suh 2015</td>
<td>-0.010</td>
<td>0.093</td>
<td>27.7</td>
<td>0.99 (0.83 to 1.19)</td>
</tr>
<tr>
<td>Velez 2015</td>
<td>-0.545</td>
<td>0.214</td>
<td>8.6</td>
<td>0.58 (0.38 to 0.88)</td>
</tr>
<tr>
<td>Yu 2015</td>
<td>-0.128</td>
<td>0.169</td>
<td>12.6</td>
<td>0.88 (0.63 to 1.22)</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td></td>
<td></td>
<td>100.0</td>
<td>0.85 (0.74 to 0.97)</td>
</tr>
<tr>
<td>Test for heterogeneity: $\chi^2=0.01, \chi^2=7.24, df=5, P=0.20, I^2=31%$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: $z=2.45, P=0.01$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sitagliptin use v no use</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wang 2014</td>
<td>0.191</td>
<td>0.080</td>
<td>64.0</td>
<td>1.21 (1.04 to 1.41)</td>
</tr>
<tr>
<td>Weir 2014</td>
<td>0.610</td>
<td>0.236</td>
<td>36.0</td>
<td>1.84 (1.16 to 2.92)</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td></td>
<td></td>
<td>100.0</td>
<td>1.41 (0.95 to 2.09)</td>
</tr>
<tr>
<td>Test for heterogeneity: $\chi^2=0.06, \chi^2=2.85, df=1, P=0.09, I^2=65%$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: $z=1.70, P=0.09$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for subgroup differences: $\chi^2=5.71, df=1, P=0.02, I^2=83%$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 6** | Risk of heart failure or hospital admission for heart failure among patients with type 2 diabetes receiving DPP-4 inhibitor treatment

<table>
<thead>
<tr>
<th>Comparison</th>
<th>No of studies (events or cases, patients)</th>
<th>DPP-4 inhibitors (events/patients)</th>
<th>Control (events or cases, patients)</th>
<th>Effect estimate (95% CI)</th>
<th>Cardiovascular morbidity at baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heart failure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Randomised controlled trials</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DPP-4 inhibitors v control</td>
<td>38 (25, 28, 292)</td>
<td>42/15701</td>
<td>33/12591</td>
<td>Pooled OR 0.97 (0.61 to 1.56)</td>
<td>Typically without CVD</td>
</tr>
<tr>
<td>Observational studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DPP-4 inhibitors v SU</td>
<td>1 (11, 616)</td>
<td>8/436</td>
<td>3/153</td>
<td>Unadjusted OR 0.88 (0.22 to 3.48)</td>
<td>With or without CVD</td>
</tr>
<tr>
<td>DPP-4 inhibitors v SU</td>
<td>1 (528, 13185)</td>
<td>NR</td>
<td>NR</td>
<td>Adjusted HR 1.10 (1.04 to 1.17)</td>
<td>No history of CVD or congestive heart failure</td>
</tr>
<tr>
<td>Sitagliptin v SU</td>
<td>1 (2, 2607)</td>
<td>1/1876</td>
<td>1/733</td>
<td>Unadjusted OR 0.39 (0.02 to 6.26)</td>
<td>NR</td>
</tr>
<tr>
<td>Sitagliptin use v no use</td>
<td>1 (457, 5027)</td>
<td>—</td>
<td>—</td>
<td>Adjusted OR 0.75 (0.38 to 1.46)</td>
<td>Admission to hospital for an acute coronary syndrome event</td>
</tr>
<tr>
<td>Hospital admission for heart failure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Randomised controlled trials</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DPP-4 inhibitors v control</td>
<td>5 (1174, 37, 028)</td>
<td>622/18554</td>
<td>522/18474</td>
<td>Pooled OR 1.13 (1.00 to 1.26)</td>
<td>CVD or multiple risk factors for cardiovascular disease</td>
</tr>
<tr>
<td>Observational studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DPP-4 inhibitors v active control (pooled estimates)</td>
<td>6 (1341, 1618295)</td>
<td>—</td>
<td>—</td>
<td>Pooled adjusted OR 0.85 (0.74 to 0.97)</td>
<td>With or without CVD</td>
</tr>
<tr>
<td>DPP-4 inhibitors v SU</td>
<td>1 (1875, 657596)</td>
<td>380/202292</td>
<td>1495/455304</td>
<td>Adjusted HR 0.84 (0.74 to 0.96)</td>
<td>With or without CVD</td>
</tr>
<tr>
<td>DPP-4 inhibitors v pioglitazone</td>
<td>2 (1060, 1031432)</td>
<td>796/776449</td>
<td>264/254983</td>
<td>Adjusted HR 0.67 (0.57 to 0.78)</td>
<td>With or without CVD</td>
</tr>
<tr>
<td>DPP-4 inhibitors v other OADs</td>
<td>1 (1118, 18744)</td>
<td>—</td>
<td>—</td>
<td>Adjusted OR 0.88 (0.63 to 1.22)</td>
<td>With or without CVD</td>
</tr>
<tr>
<td>DPP-4 inhibitors v control</td>
<td>1 (127, 3987)</td>
<td>NR</td>
<td>NR</td>
<td>Adjusted HR 0.58 (0.38, 0.88)</td>
<td>With or without CVD</td>
</tr>
<tr>
<td>Sitagliptin use v no use (pooled estimates)</td>
<td>2 (1438, 25, 638)</td>
<td>—</td>
<td>—</td>
<td>Pooled adjusted OR 1.41 (0.95 to 2.09)</td>
<td>—</td>
</tr>
<tr>
<td>Sitagliptin use v no use</td>
<td>1 (614, 16, 576)</td>
<td>339/8288</td>
<td>275/8288</td>
<td>Adjusted HR 1.21 (1.04 to 1.42)</td>
<td>With or without CVD</td>
</tr>
<tr>
<td>Sitagliptin use v no use</td>
<td>1 (824, 9062)</td>
<td>—</td>
<td>—</td>
<td>Adjusted OR 1.84 (1.16 to 2.92)</td>
<td>Heart failure at baseline</td>
</tr>
</tbody>
</table>

CVD=cardiovascular disease; SU=sulfonylurea; OR=odds ratio; HR=hazard ratio; NR=not reported; OADs=oral antidiabetic drugs.

*Nested case-control study.*
enrolled no less than 100 patients and followed up patients for 24 weeks; the third included 37 trials for analysis; the fourth included trials and observational studies. All the last three studies found that DPP-4 inhibitors were statistically associated with an increased risk of heart failure (risk ratio 1.16 (1.01 to 1.33), odds ratio 1.19 (1.03 to 1.37), odds ratio 1.15 (1.02 to 1.29), respectively).

Compared with these studies, our review has added substantial information. Firstly, we separately addressed heart failure and hospital admission as a result of heart failure. Secondly, we included both observational studies and randomised controlled trials. With respect to the trials, two important large trials were published subsequent to the previous reviews and allowed us to analyse the effect of DPP-4 inhibitors on hospital admission for heart failure. We also included additional large observational studies that carry important information regarding the risk of heart failure or admission for heart failure.

Our findings regarding the effect of DPP-4 inhibitors on heart failure were not consistent with previous meta-analyses. This difference is probably due to the fact that the previous studies were dominated by large trials reporting positive association with hospital admission for heart failure (eg, SAVOR TIMI-53), and more recent trials that have failed to find an effect were not considered.

We also found all four meta-analyses in our study to have several methodological issues. Firstly, these reviews have pooled data for heart failure and hospital admissions for heart failure. We believe that a more appropriate analysis should consider the two outcomes separately. We identified varying results when analysing the two outcomes separately. More importantly, the pooling of the two outcomes together would probably result in misleading effect estimates, when the authors aimed to assess the effect of DPP-4 inhibitors on the risk of heart failure. Another meta-analysis investigated DPP-4 inhibitors on the risk of new onset of heart failure, but this study included trials, such as SAVOR TIMI-53 and EXAMINE that already included patients with heart failure at baseline. The third meta-analysis failed to include outcome data published in ClinicalTrials.gov. The final meta-analysis combined randomised controlled trials and observational studies to generate grand effect estimates. Because of the substantial differences in the design and analysis of the type of studies, and the considerable variation in observational studies, the grand pooling will introduce misleading findings.

Implications for practice

The current evidence suggests a possible increased risk of hospital admission for heart failure in those patients with type 2 diabetes treated with DPP-4 inhibitors and with cardiovascular diseases or multiple risk factors for vascular diseases at baseline. Although the effect is small if it exists, and the associated confidence interval includes no effect, our results suggest the advisability of caution in the use of DPP-4 inhibitors for patients with type 2 diabetes who are at high risk for heart failure.

Conclusions

The relative effect of DPP-4 inhibitors on heart failure remains uncertain in patients with type 2 diabetes, given the relatively short follow-up and low quality of evidence. The current evidence suggests a small increase in the risk of hospital admission for heart failure in patients with existing cardiovascular diseases or multiple risk factors for vascular diseases. Additional randomised controlled trials enrolling patients with existing cardiovascular diseases or multiple risk factors for vascular diseases will be required to definitively assess the effect of DPP-4 inhibitors on such patients. Such trials, if enrolling patients at high risk of exacerbation and admission, may be feasible. In the meantime, the possible increase in hospital admission for heart failure could be one issue that patients and clinicians consider in choosing antidiabetic drug treatment for patients with existing cardiovascular diseases.

AUTHOR AFFILIATIONS

1Chinese Evidence-based Medicine Centre, West China Hospital, Sichuan University, Chengdu 610041, Sichuan, China
2Department of Endocrinology and Metabolism, West China Hospital, Chengdu University
3West China School of Pharmacy, Sichuan University, Chengdu University
4Norwegian Knowledge Centre for the Health Services, Oslo, Norway
5Department of Medicine, Inlandet Hospital Trust, Gjøvik, Norway
6Department of Hygiene and Dietetics, Jagiellonian University Medical College, Krakow, Poland
7Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, ON, Canada
8Faculty of Medicine, University of Toronto, Toronto, ON, Canada
9Department of Medicine, University of British Columbia, Vancouver, BC, Canada
10Department of Anesthesia, McMaster University, Hamilton
11Michael G DeGroote Institute for Pain Research and Care, McMaster University, Hamilton
12Stanford Prevention Research Center, Department of Medicine, Stanford University, Stanford, CA, USA
13Department of Anaesthesia and Pain Medicine, Hospital for Sick Children, Toronto, ON, Canada
14Department of Medicine, Universidad Peruana Cayetano Heredia, Lima, Peru
15Department of Medical Administration, 363 Hospital, Chengdu, Sichuan, China
16Second Hospital of Lanzhou University, Lanzhou, Gansu, China
17Department of Medicine, McMaster University, Hamilton
18Department of Medicine, McMaster University, Hamilton
19Department of Medicine, McMaster University, Hamilton
20We thank Daphne Plaut for developing the search strategy and conducting the initial literature search.

Contributors: XS and SL conceived the study. XS acquired the funding. XS and LL had full access to all of the data in the study, and take responsibility for the integrity of the data and the accuracy of the data analysis. XS and LL designed the study. XS and LL developed and tested the data collection forms. LL, KD, JL, PZ, LZ, JS, MMB, ZNS, EW, JWB, SE, GM, LPR, POY, YW, QC, and XS acquired the data. LL and KD took responsibility for the integrity of the data and the accuracy of the data analysis. XS and SL designed the study. XS and LL developed and tested the data collection forms. LL, KD, JL, PZ, LZ, JS, MMB, ZNS, EW, JWB, SE, GM, LPR, KD, JL, PZ, LZ, JS, YW, and QC critically revised the manuscript. XS is the guarantor.

Funding: This study was supported by the National Natural Science Foundation of China (grant no 71573183), “Thousand Youth Talents Plan” of China (grant no D102400J) and Sichuan Province, and Young
investigator Award of Sichuan University (grant no. 2013JQ04A37). These funders had no role in the study design, writing of the manuscript, or decision to submit this or future manuscripts for publication. SL is funded by the National Natural Science Foundation of China (grant No 81400811 and 21510008). JNS is funded by the Canadian Diabetes Association. JB is funded by a New Investigator Award from the Canadian Institutes of Health Research and Canadian Chiropractic Research Foundation. SE is funded by MITACS Elevate and Restrocomp Postdoctoral Awards.

**Competing interests:** All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coiDisclosure.pdf and declare: support for the submitted work in the previous three years, and no other relationships or activities that could appear to have influenced the submitted work.

**Ethical approval:** Not required.

**Data sharing:** No additional data available.

The lead author (the manuscript’s guarantor) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained. This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) 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/3.0/
49
46
44
40
37
36
35
30
29
28
27
26
25
24
23
22
21
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1


© BMJ Publishing Group Ltd 2016

Web appendix 1: Search strategies
Web appendix 2: Risk of bias of included randomised controlled trials
Web appendix 3: Supplementary forest plots
Web appendix 4: Risk of bias of included cohort studies
Web appendix 5: Risk of bias of included case-control studies