RQ 1.2: Subsequent pharmacological therapy for the management of type 2 diabetes.

August 2024

Beatrice C Downing and Hugo Pedder, Technical Support Unit, University of Bristol

Research Question

This analysis focused on subsequent therapies for management of type 2 diabetes. In this context, study participants would be receiving the trialled treatments against a background of initial management, which could include treatments included in the subsequent treatment set. Study populations varied in their existing treatments and in the use of rescue therapies.

Methods

Network meta-analyses were conducted in WinBUGS, version 1.4.3, using standard TSD code (Dias 2011), adapted for the data reporting style (WinBUGS code).

Normal likelihood with identity link function for estimation of mean change, additive treatment effects

Used for the analysis of change in HbA1c and the sensitivity analysis of change in weight (kg).

Models were run using TSD code TSD2-8 (Normal shared), which estimates treatment effects jointly, updated to allow for direct use of mean differences reported in different formats: as arm- or contrast-level change-from-baseline (CFB) and arm- or contrast-level mean value at follow-up.

Normal likelihood with log link function for estimation of mean change, proportional treatment effects

This model was used for the analysis of change in weight (kg).

Models were run using TSD code TSD2-5 (Normal id), updated to allow for direct use of mean differences reported in different formats: as change from baseline or mean value at follow-up. Within this model, it was necessary to assume correlation between baseline and follow-up measures to combine relative effects from studies reporting the change from baseline and studies reporting relative effects as baseline and follow-up values. A value of 0.5 was used for the correlation to calculate the SE around the change from baseline, which is somewhat conservative (Balk et al. 2012, Daly et al. 2021).

In this model, studies not reporting baseline values and studies reporting contrast-level estimates were excluded from the dataset for analysis.

Shared parameter model for estimation of hazard ratios

This model jointly modelled binomial data on the number of events on each arm at a given follow-up time, and continuous data on the hazard ratio between arms. Binary data were modelled with a clog-

log link to include exposure time, following the code in TSD2-2 (binomial likelihood, clog-log link, Dias 2011). This assumes a constant underlying event rate (hazards are exponentially distributed) and informs a log-hazard ratio as the treatment effect. Continuous data – in the form of log-hazard ratios – were modelled following the code in TSD2-7 (Normal difference).

This model was used in the estimation of treatment effects for cardiovascular mortality, hospitalisation for heart failure, three-point MACE, non-fatal stroke, non-fatal myocardial infarction, end-stage kidney disease and unstable angina.

Where studies reported both a hazard ratio (with its confidence interval) and the number of events, the hazard ratio was preferred for synthesis in the NMA. Studies reporting zero events on all arms were excluded from the analysis because they provided no information on relative effects.

Model selection

The choice of whether to proceed with the NMA with fixed-effect (FE) or random-effects (RE) structure on the treatment differences was made by first considering the network structure, then by comparing model fit statistics between the two models, taking into account the size and precision of the between-study variation estimated under the RE model.

Where all edges in the network were informed only by a single study, only FE models were fitted because there was insufficient evidence from which to estimate the between-study variation. Where the randomised trial evidence formed star networks, there would be no indirect evidence (appearing as loops in the network) from which to estimate inconsistency. Therefore, inconsistency – or unrelated mean effect – models were not run on star networks.

Where DIC in the FE and RE models differed by >3 units, the model with the lower DIC was chosen. Where DIC values in FE and RE models were within 2 units, the FE model was selected provided that total residual deviance was similar to the number of data points and either the median of between-study SD was low or between-study SD was poorly estimated from the data (skewed distribution and a large uncertainty interval). There were no cases where differences in FE DIC and RE DIC were within 2 units and the between-study SD was both high and precisely estimated.

Assessing inconsistency with Unrelated Mean Effect models

NMA assumes that the included studies are similar in terms of factors that might interact with the intervention effects (effect modifiers). So, the relative effect of intervention B vs intervention A would be expected to be similar in all of the studies (if they had included A and B interventions). We can assess this assumption by measuring statistical heterogeneity, and also by checking if the direct and indirect estimates are in agreement when there are loops of evidence in the network. To determine if there is evidence of inconsistency, the selected consistency model (fixed or random effects) was compared to an “inconsistency”, or unrelated mean effects (UME), model (Dias 2013). The latter is equivalent to having separate, unrelated, meta-analyses for every pairwise contrast, with a common variance parameter assumed in the case of random effects models. We further explored inconsistency using dev-dev plots, which can highlight individual studies that are contributing to inconsistency. These study arms appear in the area shaded grey in the included dev-dev plots.

Imprecise estimates, where events were rare

Treatment effects (log-hazard ratios) were not estimable for all treatments for all outcomes. In these cases, the small number of events mean that there was insufficient evidence to return a suitably precise estimate and the estimated log-hazard ratio from NMA only (or largely) reflects the vague

prior in the model. Where this was the case, the treatment effect is reported as 'Not estimable'. This affected five outcomes in the population with diabetes and chronic kidney disease (CKD) and two outcomes in the population with diabetes and atherosclerotic cardiovascular disease (AS-CVD).

Population with CKD:

- Cardiovascular mortality: ertugliflozin, linagliptin, liraglutide, semaglutide (oral delivery)
- Hospitalisation for heart failure: ertugliflozin, semaglutide (oral delivery)
- Non-fatal stroke: linagliptin, liraglutide, dapagliflozin
- Non-fatal myocardial infarction: linagliptin, liraglutide and sitagliptin
- End-stage kidney disease: saxagliptin (very wide interval on log-hazard scale)

Population with CVD:

- Non-fatal stroke: glimepiride
- Non-fatal myocardial infarction: vildagliptin

Definition of the subpopulation with chronic kidney disease (CKD)

These analyses focus on people with type 2 diabetes mellitus (T2D) and chronic kidney disease (CKD). Studies in people with CKD primarily were not included in the evidence, including where a T2D subgroup was reported.

Summary

In this subpopulation, there was randomised controlled trial (RCT) evidence for eight outcomes:

- Change in HbA1c
- Change in body weight
- Hospitalisation for heart failure
- Cardiovascular mortality
- Three-point MACE
- Non-fatal stroke
- Non-fatal myocardial infarction
- End-stage kidney disease

Fixed-effect models were preferred on model fit (total residual deviance and DIC) for all eight outcomes; therefore, none of the NMAs presented for this subpopulation model the variation between studies making the same treatment comparison.

Most networks in this subpopulation were star networks, without loops providing indirect evidence. Therefore, inconsistency between direct and indirect evidence could only be assessed for change in HbA1c.

Subpopulation with CKD: Change in HbA1c

This analysis included 24 studies of 18 treatments (Figure 1). NMA with fixed-effect structure on treatment effects was sufficient for modelling this outcome on the basis of model fit (Table 25).

There is evidence supporting reduction in % HbA1c compared to placebo for pioglitazone, glimepiride and all represented members of the SGLT2-inhibitor class. There was clear evidence supporting a reduction in % HbA1c for dulaglutide, semaglutide by both routes and liraglutide (GLP1-receptor agonists) and linagliptin and saxagliptin (DPP4 inhibitors) (Figure 2). The evidence for a reduction in % HbA1c relative to placebo was weaker for insulin, exenatide, vildagliptin and sitagliptin, where there was greater uncertainty around the effect of treatment.

Semaglutide by oral and subcutaneous routes showed the largest precisely estimated reduction in % HbA1c, whilst glimepiride's median estimated effect was large (Table 1), there was considerable uncertainty around the strength of the reduction in % HbA1c. There was evidence for a reduction in % HbA1c using oral semaglutide relative to canagliflozin, dapagliflozin, dapagliflozin with saxagliptin, empagliflozin, ertugliflozin, linagliptin, sitagliptin, vildagliptin, dulaglutide, exenatide, liraglutide and insulin. There was evidence for a reduction in % HbA1c using subcutaneous semaglutide relative to canagliflozin, dapagliflozin, ertugliflozin, dulaglutide, sitagliptin, vildagliptin and insulin. For the full table of active-active comparisons, see tab 'Treatment Direct Effects' in RQ1.2 results file KD HbA1c.

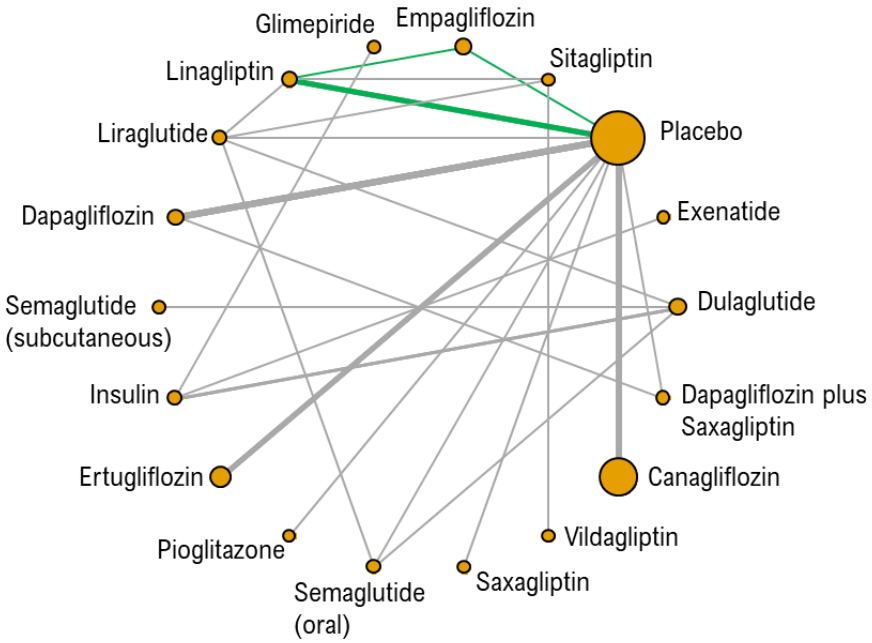

Figure 1. Network of evidence for change in HbA1c in the subpopulation with CKD. The loop of evidence containing inconsistency is shown in green.

Table 1. Change in HbA1c relative to placebo for all 17 active treatments in the network, with random-effects structure on treatment effects.

Class	Treatment	Change in HbA1c Posterior median (95%CrI)
DPP-4i	Linagliptin	-0.54 (-0.66, -0.41)
	Saxagliptin	-0.63 (-1.24, -0.02)
	Sitagliptin	-0.21 (-0.58, 0.15)
	Vildagliptin	-0.19 (-0.70, 0.31)
GLP-1RA	Dulaglutide	-0.42 (-0.72, -0.12)
	Exenatide	-0.09 (-0.85, 0.66)
	Liraglutide	-0.54 (-0.71, -0.37)
	Semaglutide (Oral)	-0.92 (-1.15, -0.70)
	Semaglutide (Subcutaneous)	-0.82 (-1.19, -0.44)
Insulin	Insulin	-0.32 (-0.71, 0.06)
SGLT2 with DPP-4i	Dapagliflozin with Saxagliptin	-0.58 (-0.79, -0.36)
SGLT-2i	Canagliflozin	-0.20 (-0.32, -0.08)
	Dapagliflozin	-0.23 (-0.39, -0.06)
	Empagliflozin	-0.51 (-0.63, -0.39)
	Ertugliflozin	-0.10 (-0.22, 0.02)
Sulfonylurea	Glimepiride	-0.92 (-1.70, -0.13)
TZD	Pioglitazone	-0.81 (-1.47, -0.16)

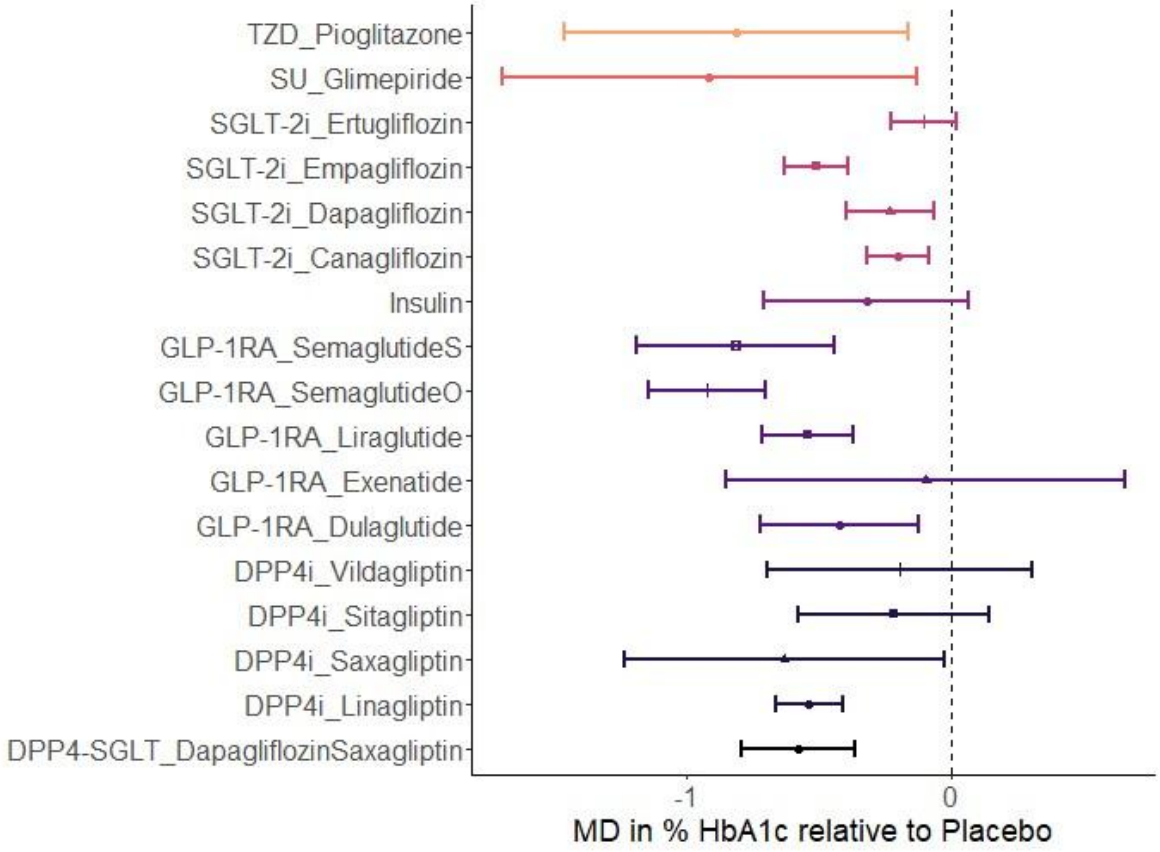

Figure 2. Mean difference (MD) in % HbA1c relative to placebo in the subpopulation with diabetes and CKD. Mean treatment effects are shown as points, with the 95% credible interval. Line colour corresponds to treatment class, with point shape distinguishing different treatments within the same class.

Inconsistency in network evidence

We note inconsistency on the loop of evidence between placebo, empagliflozin and linagliptin. This was linked to high deviance for the study Raman 2022, which trialled empagliflozin against linagliptin. Raman et al. 2022 reported empagliflozin to be more effective by 0.3 units; however, indirect evidence from the network suggests that linagliptin and empagliflozin are similarly effective. The indirect treatment difference for linagliptin vs empagliflozin was estimated from the direct evidence of each of these treatments compared with placebo (Table 2). Studies comprising the loop were inspected for discrepancies in inclusion/exclusion criteria that may have led to effect modification, but no clear differences could be identified.

Table 2. Inconsistency in evidence for change in HbA1c within the subpopulation with CKD.

Comparison	Source	Mean treatment difference (95% CrI)

Empa vs Placebo	*Barnett 2014*	*-0.46 (-0.58, -0.33)*
	Synthesis of direct evidence	**-0.46 (-0.58, -0.33)**
	NMA, direct and indirect	**-0.51 (-0.63, -0.39)**
Lina vs Placebo	*Moeinzadeh 2021*	*-0.14 (-0.71, 0.43)*
	Groop 2017	*-0.60 (-0.78, -0.43)*
	McGill 2013	*-0.72 (-1.03, -0.41)*
	Synthesis of direct evidence	**-0.60 (-0.74, -0.45)**
	NMA, direct and indirect	**-0.54 (-0.66, -0.41)**
Lina vs Empa	*Raman 2022*	*0.34 (0.01, 0.67)*
	Synthesis of direct evidence	**0.34 (0.01, 0.67)**
	NMA, direct and indirect	**-0.03 (-0.19, 0.14)**

Subpopulation with CKD: Change in weight

This analysis included six studies of six treatments (Figure 3). NMA with fixed-effects structure on treatment effects was sufficient for modelling this outcome on the basis of model fit (Table 26), which is to be expected given that only one treatment comparison was observed in more than one study.

Treatment effects were very precisely estimated and provide clear evidence of a very small reduction in weight for those patients on the SGLT2-inhibitors in the network: canagliflozin, dapagliflozin, empagliflozin and ertugliflozin (Table 3). The evidence for linagliptin was more uncertain, with the 95% credibility interval supporting the probability of weight-loss, weight-neutral and weight-gain effects. No clear differences were noted between active treatments (Figure 5).

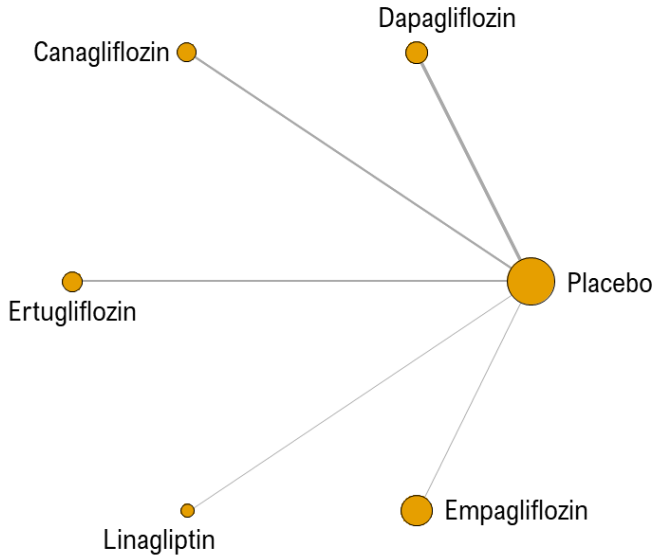

Figure 3. Network of evidence for proportional change in weight in the subpopulation of people with diabetes and CKD.

Table 3. Predicted treatment effects relative to placebo for all active treatments in the network. This was a proportional model, therefore treatment effects are estimated as the ratio of means. Mean differences in kg are predicted for patients with a baseline weight of 88.3 and a mean CFB on the placebo treatment of -1.92kg, as reported by Fioretto 2018 for the placebo arm.

Class	Treatment	Ratio of change (95% CrI)	Mean difference in kg (95% CrI)
SGLT-2i	Canagliflozin	0.99 (0.98, 1.00)	-1.02 (-1.82, -0.20)
	Dapagliflozin	0.98 (0.97, 0.99)	-1.50 (-2.23, -0.76)
	Empagliflozin	0.98 (0.98, 0.99)	-1.69 (-2.18, -1.19)
	Ertugliflozin	0.98 (0.97, 0.99)	-1.70 (-2.36, -1.04)
DPP-4i	Linagliptin	0.99 (0.92, 1.05)	-1.31 (-6.88, 4.65)

Sensitivity analysis: analysis of weight change as an additive measure

Analysis of weight change as a proportional measure required that included studies report both a) weight at baseline and b) arm-level outcomes. Enforcing these requirements meant excluding the majority of studies reporting weight change. Therefore, a sensitivity analysis was run within which weight change was modelled as additive, meaning that the effect of treatment, in kg, was the same regardless of the participant's baseline weight. Whilst there was more evidence available for the sensitivity analysis, the committee confirmed that they would expect weight change to show proportional change to treatment, therefore this was the base-case analysis.

This analysis included 18 studies of 14 treatments, with the fixed-effect structure on treatment effects being adequate for modelling purposes.

For the treatments included in the base-case analysis (weight change as a proportional measure), the treatment effects are consistent in the sensitivity analysis (Table 4), with evidence to support weight loss with use of the included SGLT-2 inhibitors (canagliflozin, dapagliflozin, empagliflozin and

ertugliflozin), and support for weight-neutral effect (with a trend for weight reduction) for linagliptin relative to placebo (Figure 5).

Turning to those treatments where there was only evidence for weight change as an additive measure, there was strong evidence to support reduction in weight for those receiving liraglutide and for those receiving semaglutide (both oral and subcutaneous delivery) relative to those on the placebo arm (Table 4). The evidence for dulaglutide and exenatide was uncertain, with the possibility of weight-reducing, weight-neutral and weight-increasing effects within the wide credible intervals. The treatment effect of glimepiride was highly uncertain, with the 95% credible interval spanning from a reduction in weight of 8.4kg to an increase in weight of 9.3kg. There was evidence of an increase in weight in those receiving insulin, relative to placebo.

For a full comparison of all active treatments, see the 'Treatment Direct Effects' tabs in RQ1.2 results files KD kg and KD kg (additive).

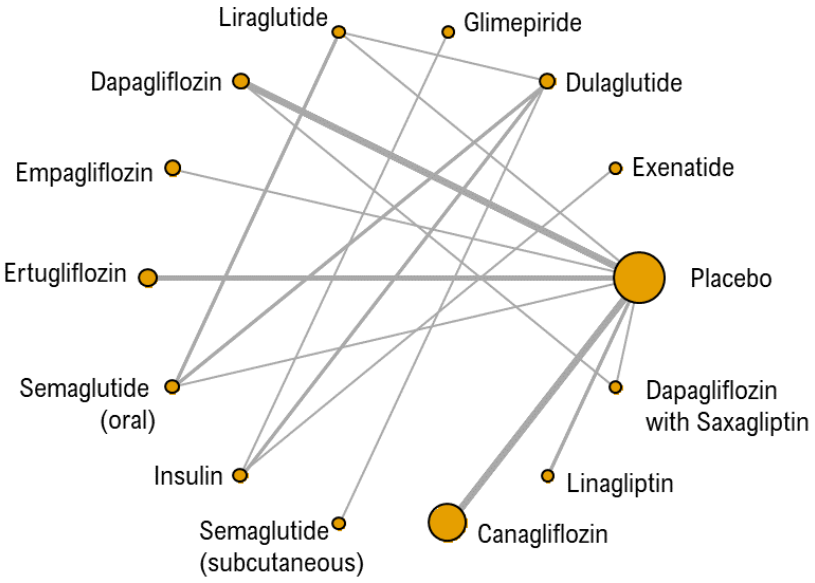

Figure 4. Network of evidence for additive change in weight in the subpopulation of people with diabetes and CKD.

Table 4. Change in weight relative to placebo as an additive measure in the subpopulation of people with diabetes and CKD. Treatment effects are presented as mean differences (in kg) with the posterior median ranking with its range.

Class	Treatment	Mean difference in kg Median (95% CrI)	Rank Median (95% CrI)
GLP-1RA	Dulaglutide	-0.28 (-1.25, 0.70)	10 (7, 12)
	Exenatide	0.71 (-1.62, 3.05)	12 (5, 13)
	Liraglutide	-1.09 (-1.82, -0.36)	7 (4, 10)
	Semaglutide (oral)	-2.68 (-3.37, -1.99)	2 (1, 4)
	Semaglutide (subcutaneous)	-2.79 (-4.07, -1.47)	2 (1, 5)
Sulfonylurea	Glimepiride	0.41 (-8.43, 9.29)	12 (1, 14)
Insulin	Insulin	3.39 (1.89, 4.89)	14 (13, 14)
SGLT-2i	Canagliflozin	-0.90 (-1.38, -0.43)	8 (5, 10)

	Dapagliflozin	-1.41 (-2.07, -0.76)	6 (3, 9)
	Empagliflozin	-1.55 (-2.00, -1.10)	5 (3, 8)
	Ertugliflozin	-2.05 (-2.53, -1.58)	3 (1, 5)
SGLT-2i with DPP-4i	Dapagliflozin with Saxagliptin	-0.04 (-1.34, 1.26)	11 (6, 13)
DPP-4i	Linagliptin	-0.95 (-3.21, 1.32)	8 (1, 13)

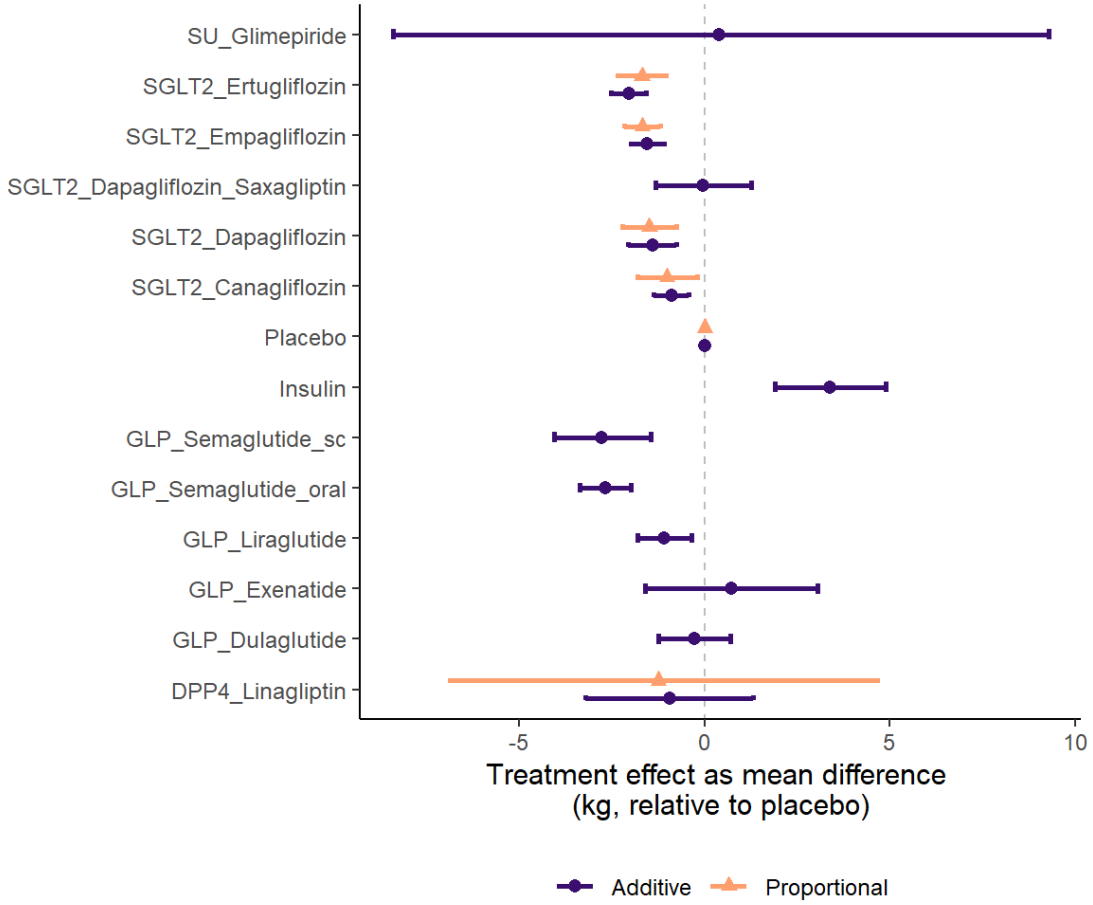

Figure 5. Mean difference in body weight (kg) on treatment, relative to placebo, in the subpopulation with diabetes and CKD. Mean treatment effects are shown as points, with the 95% credible interval. Treatment effects drawn from FE additive (purple circles) and proportional (orange triangles) models. Note that only five active treatments were represented in the proportional model.

Subpopulation with CKD: Cardiovascular (CV) mortality

This analysis included nine trials of eight treatments (Figure 6). Three studies reported hazard ratios and six reported the number of events.

Treatment effects for semaglutide (oral delivery), ertugliflozin, linagliptin and liraglutide were not estimable for this outcome, with estimates for ertugliflozin and linagliptin being particularly uncertain. The evidence for ertugliflozin was drawn from a single study of 467 people, Grunberger 2018, that trialled two doses of ertugliflozin against placebo. Zero CV deaths were reported on the placebo arm and one of the ertugliflozin arms, whilst a single CV death was reported on the second ertugliflozin arm. The evidence for linagliptin was drawn from a single study of 360 people, Groop

2017, that trialled linagliptin against placebo. Zero CV deaths were reported on the placebo arm, and two on the linagliptin arm.

There was clear evidence for a decreased hazard of CV mortality compared to placebo for empagliflozin (Table 5). There was weaker evidence for a decreased hazard of CV mortality compared to placebo for the other SGLT2-inhibitors with evidence: canagliflozin and dapagliflozin. No precisely estimated differences were noted between active treatments (for the full table of active-active comparisons, see tab 'Treatment Direct Effects' in RQ1.2 results file KD CVM).

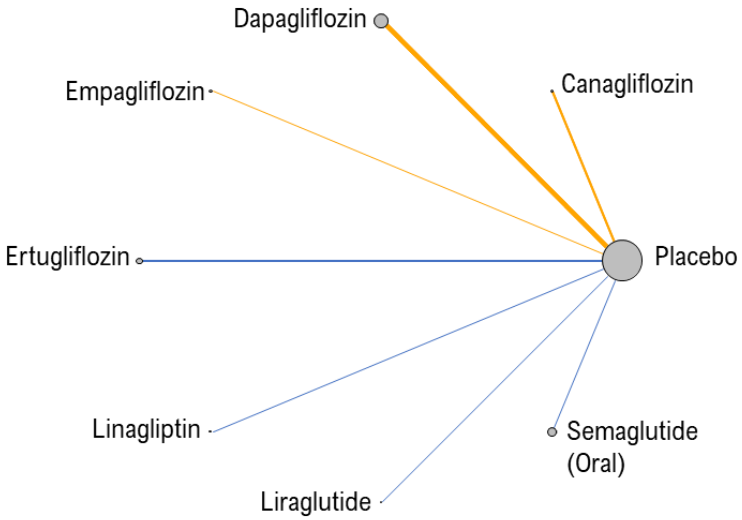

Figure 6. Network of evidence for cardiovascular mortality in the subpopulation of diabetes patients with CKD. Network edges where data were sparse (zero events on one study arm) are shown in blue.

Table 5. Hazard ratios for cardiovascular mortality in the subpopulation with diabetes and CKD, relative to placebo. Posterior medians and ranges of rankings are also presented - these are less easy to interpret where hazard ratios are uncertain.

Class	Treatment	Hazard ratio Median (95% CrI)	Rank Median (95% CrI)
SGLT-2i	Canagliflozin	0.79 (0.62, 1.01)	3 (1, 5)
	Dapagliflozin	0.90 (0.71, 1.15)	4 (2, 6)
	Empagliflozin	0.71 (0.52, 0.97)	2 (1, 5)
	Ertugliflozin	*Not estimable*	*Not estimable*
GLP-1RA	Liraglutide	*Not estimable*	*Not estimable*
	Semaglutide (oral)	*Not estimable*	*Not estimable*
DPP-4i	Linagliptin	*Not estimable*	*Not estimable*

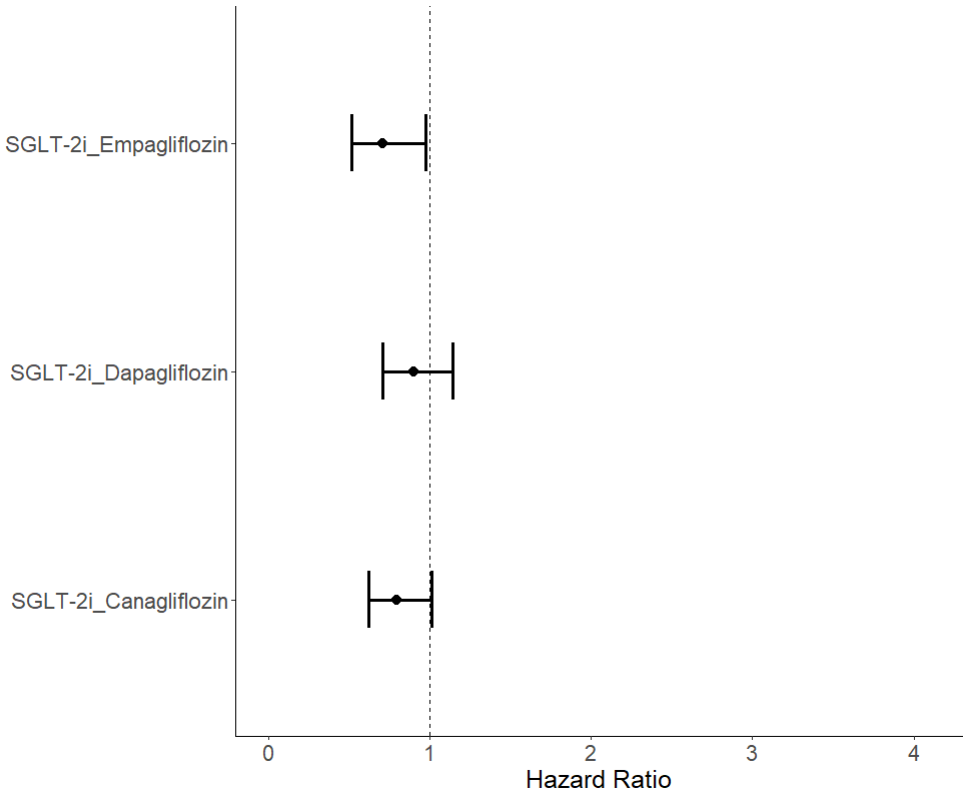

Figure 7. Hazard ratios for cardiovascular mortality, relative to placebo, in the subpopulation with T2D and CKD. Imprecisely estimated treatment effects for ertugliflozin, liraglutide, oral semaglutide and linagliptin not plotted.

Subpopulation with CKD: hospitalisation for heart failure

This analysis included nine trials of nine treatments (Figure 8). Four studies reported hazard ratios, and five reporting the number of events alone. Treatment effects for semaglutide (oral delivery) and ertugliflozin were not estimable for this outcome.

There was clear evidence for decreased hazard of hospitalisation for heart failure compared with placebo for all SGLT2-inhibitors with evidence: canagliflozin, dapagliflozin and empagliflozin. There was weaker evidence of decreased hazard of hospitalisation for heart failure relative to placebo for linagliptin. There was substantial uncertainty around the conclusion of any protective effect of sitagliptin and liraglutide.

There was evidence to support lower hazard of hospitalisation for heart failure in those given canagliflozin vs in those given linagliptin, with a median hazard ratio of 0.72 (95% CrI: 0.51, 1.01). For the full table of active-active comparisons, see tab ‘Treatment Direct Effects’ in RQ1.2 results file KD hosp.

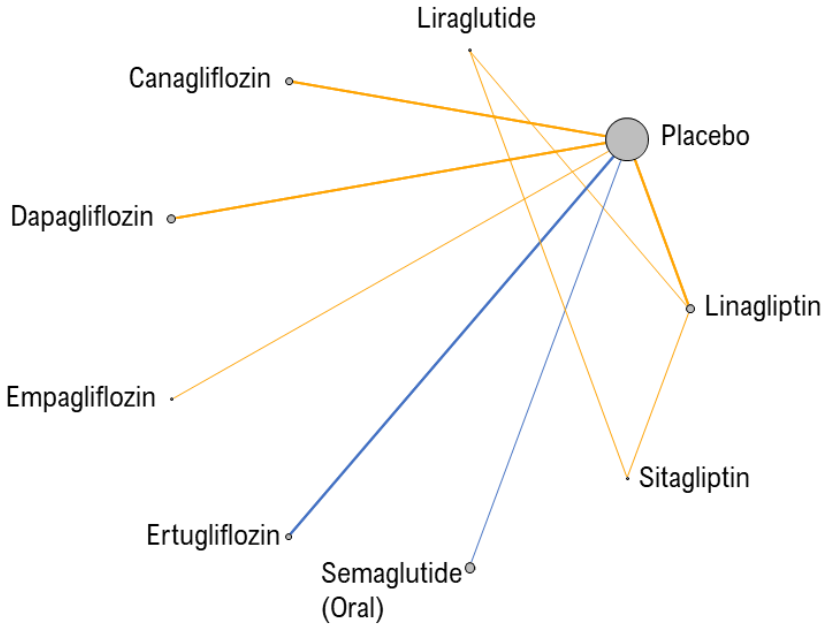

Figure 8. Network of evidence for hospitalisation for heart failure in the subpopulation of people with diabetes and CKD. Network edges where data were sparse (zero events on one study arm) are shown in blue.

Table 6. Hazard ratios for hospitalisation for HF in the subpopulation with diabetes and CKD, relative to placebo. Posterior medians and ranges of rankings are also presented - these are less easy to interpret where hazard ratios are uncertain and include treatments with non-estimable intervals in their calculation

Class	Treatment	Hazard ratio Median (95% CrI)	Rank Median (95% CrI)
SGLT-2i	Canagliflozin	0.61 (0.47, 0.79)	4 (2, 7)
	Dapagliflozin	0.72 (0.57, 0.91)	6 (3, 8)
	Empagliflozin	0.61 (0.42, 0.88)	4 (2, 7)
	Ertugliflozin	*Not estimable*	*Not estimable*
GLP-1RA	Liraglutide	0.82 (0.14, 4.74)	7 (2, 9)
	Semaglutide (oral)	*Not estimable*	*Not estimable*
DPP-4i	Linagliptin	0.84 (0.68, 1.04)	7 (5, 8)
	Sitagliptin	0.75 (0.13, 4.48)	6 (2, 9)

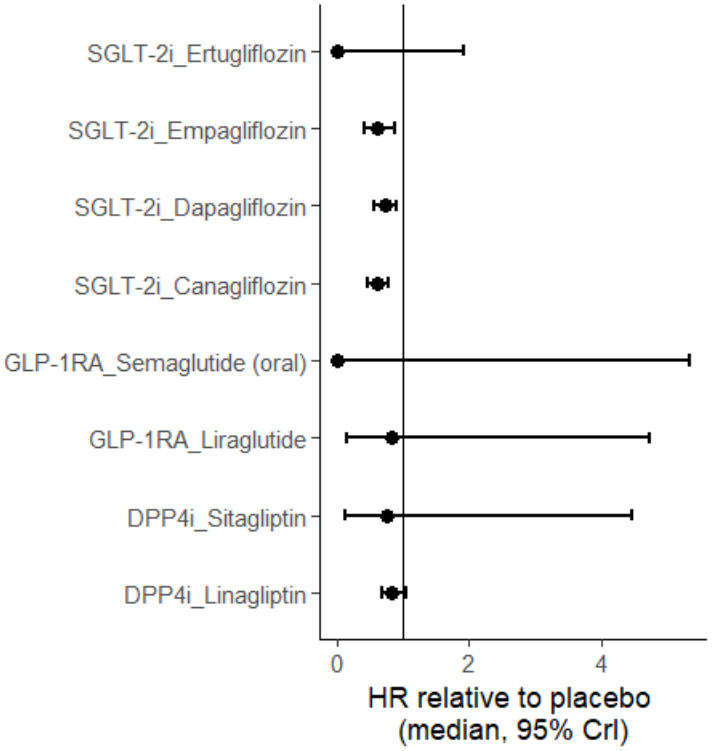

Figure 9. Hazard ratio for hospitalisation for heart failure, relative to placebo, in the population with T2D and CKD.

Subpopulation with CKD: three-point MACE

This analysis included three trials of three treatments (Figure 10). Two studies reported the hazard ratio and the number of events, with one reporting the number of events alone. Between 3% and 12% of study participants experienced an event included in the composite three-point MACE.

There was clear evidence suggesting lower hazard of three-point MACE for canagliflozin vs placebo, and weaker evidence for a lower hazard for dapagliflozin vs placebo (Table 7). There was no clear difference in hazard rates when comparing canagliflozin and dapagliflozin (for the full table of active-active comparisons, see tab 'Treatment Direct Effects' in RQ1.2 results file KD MACE3).

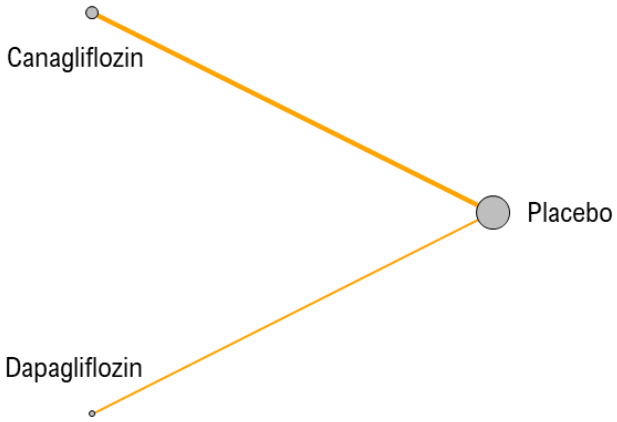

Figure 10. Network of evidence for three-point MACE in the subpopulation of people with diabetes and CKD.

Table 7. Hazard ratios and median rank with 95% CrI for three-point MACE in the subpopulation with diabetes and CKD, relative to placebo.

Class	Treatment	Hazard ratio Median (95% CrI)	Rank Median (95% CrI)
SGLT-2i	Canagliflozin	0.82 (0.69, 0.97)	1 (1, 2)
	Dapagliflozin	0.93 (0.80, 1.08)	2 (1, 3)

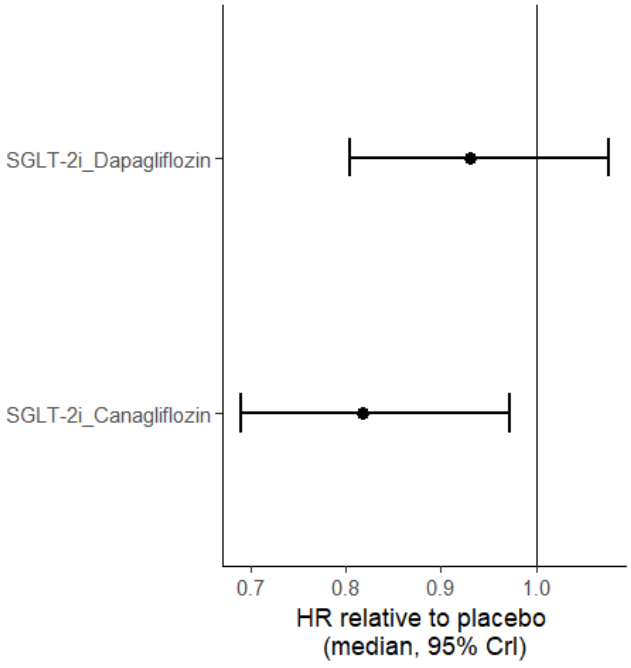

Figure 11. Hazard ratio for three-point MACE, relative to placebo, in the population with T2D and CKD.

Subpopulation with CKD: non-fatal stroke

This analysis included three two-armed trials of four treatments (Figure 12). Random-effect model structures were not considered for non-fatal stroke because there was insufficient evidence to estimate the between-study variability.

Non-fatal stroke was rare in all three trials:

- Fioretto 2018 reported one event on the placebo arm in a trial of dapagliflozin vs placebo in 321 people
- Davies 2016 reported on event on the liraglutide arm in a trial of liraglutide vs placebo in 277 people
- McGill 2013 reported one event on each arm, in a trial of linagliptin vs placebo in 133 people.

Whilst the NMA generated effect estimates for these events, these largely reflect the wide prior uncertainty in treatment effect, with an unfeasibly high upper bound to the 95% credible interval for linagliptin (Table 8), an unfeasibly low median for dapagliflozin (Table 8) and an unfeasibly large 95% credible interval for liraglutide (Figure 13).

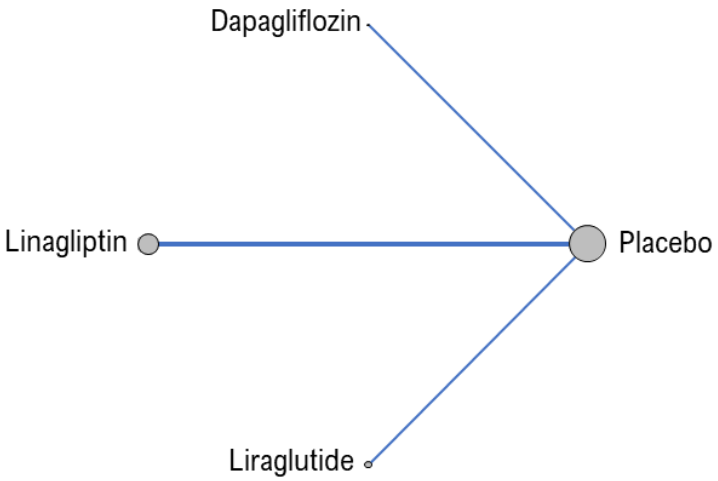

Figure 12. Network of evidence for non-fatal stroke in the subpopulation of people with diabetes and CKD. Network edges where data were sparse (e.g. zero or single events on one study arm) are shown in blue.

Table 8. Hazard ratios and median rank with 95% CrI for non-fatal stroke in the subpopulation with diabetes and CKD, relative to placebo.

Class	Treatment	Hazard ratio Median (95% CrI)
SGLT-2i	Dapagliflozin	<0.01 (<0.01, 4.30)
DPP4-i	Linagliptin	0.84 (0.01, 28.75)
GLP-1RA	Liraglutide	*Not estimable*

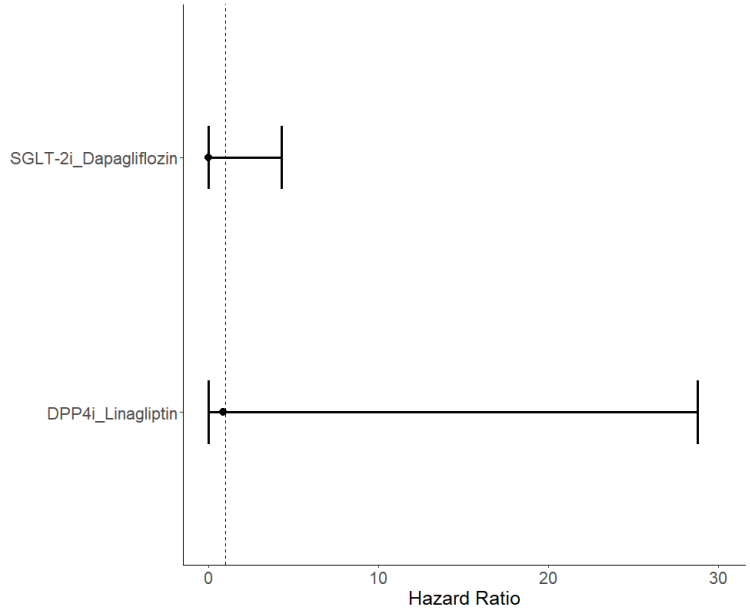

Figure 13. Hazard ratio for non-fatal stroke, relative to placebo, in the population with T2D and CKD. The effect estimate for liraglutide could not be precisely estimated, so is not plotted.

Subpopulation with CKD: non-fatal myocardial infarction (MI)

This analysis included three trials of four treatments (Figure 14). Small numbers of non-fatal MI were observed in all three trials:

- Groop 2017 reported one event on the linagliptin arm, in a trial of linagliptin vs placebo in 360 people.
- McGill 2013 reported six events over both arms, in a trial of linagliptin (four events) against placebo (two events) in 133 people.
- Hiramatsu 2018 reported six events over three arms, in a trial of linagliptin (two events), liraglutide (one event) and sitagliptin (three events) in 139 people.

All effect estimates were extremely uncertain (Table 9); there was no strong evidence that the rate of non-fatal MI was higher on trial arms where patients received linagliptin, sitagliptin or liraglutide than on the placebo arm. No clear differences were noted between active treatments (for the table of active-active comparisons, see tab 'Treatment Direct Effects' in RQ1.2 results file KD MI).

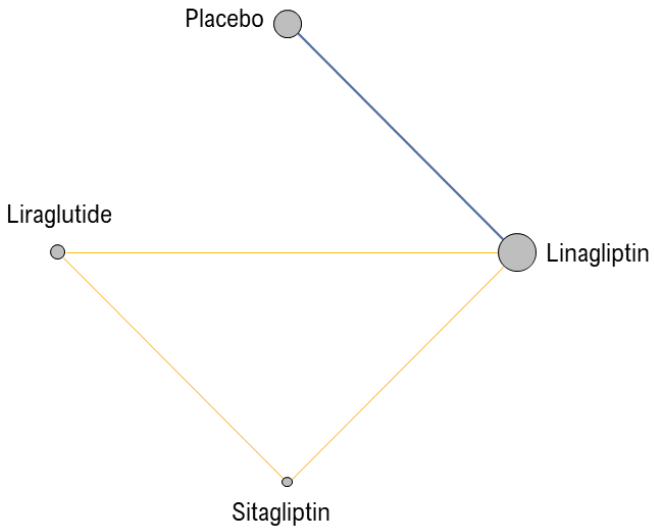

Figure 14. Network of evidence for non-fatal MI in the subpopulation of people with diabetes and CKD. Network edges where data were sparse (zero events on one study arm) are shown in blue.

Table 9. Hazard ratios and median rank with 95% CrI for non-fatal MI in the subpopulation with diabetes and CKD, relative to placebo.

Class	Treatment	Hazard ratio Median (95% CrI)	Rank Median (95% CrI)
Placebo	Placebo	*Reference*	2 (1, 4)
DPP4-i	Linagliptin	2.39 (0.50, 16.38)	3 (1, 4)
	Sitagliptin	3.45 (0.29, 54.17)	4 (1, 4)
GLP-1RA	Liraglutide	0.90 (0.01, 20.14)	2 (1, 4)

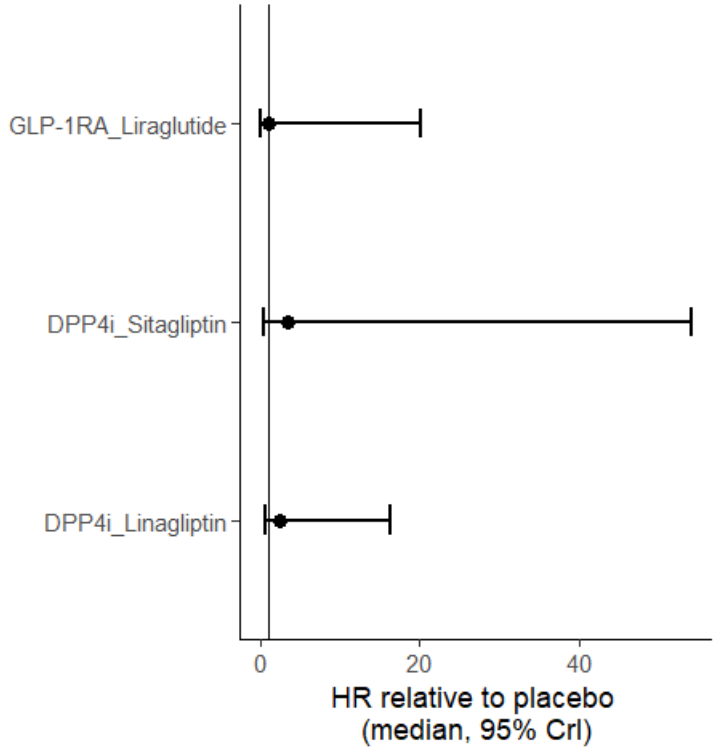

Figure 15. Hazard ratio for non-fatal MI, relative to placebo, in the population with T2D and CKD.

Subpopulation with CKD: end-stage kidney disease (ESKD)

This analysis included three trials of four treatments (Figure 16). Two of these trials were small, with small numbers of patients developing ESKD, however the rate of ESKD development was relatively similar between trials with 1-2% of participants developing ESKD in Kohan 2014 and Nowicki 2011A, and 6% of participants developing ESKD in Perkovic 2019.

- Kohan 2014 reported four events across three study arms: one on the placebo arm (1/85 participants), one of the first dapagliflozin arm (1/83) and two on the second dapagliflozin arm (2/85).
- Nowicki 2011A reported two events on the placebo arm, when trialling saxagliptin (0/85) against placebo (2/85).
- Perkovic 2019 reported 281 events across two study arms and a hazard ratio of 0.68 (95% CI: 0.54, 0.86) for canagliflozin (116/2202) vs placebo (165/2199).

There was clear evidence suggesting lower hazard of ESKD on canagliflozin vs placebo (Table 10). There was also evidence for a lower hazard of ESKD with saxagliptin vs placebo, though this is based on the two events observed on the placebo arm of Nowicki et al 2011. No precisely estimated differences were noted between active treatments (for the full table of active-active comparisons, see tab 'Treatment Direct Effects' in RQ1.2 results file KD ESKD). Random-effect model structures were not considered for end-stage kidney disease because there was insufficient evidence to estimate the between-study variability.

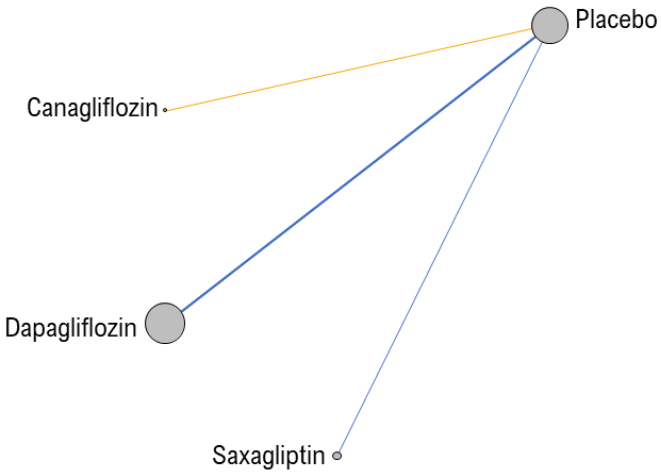

Figure 16. Network of evidence for end-stage kidney disease in the subpopulation of people with diabetes and CKD. Network edges where data were sparse (zero events on one study arm) are shown in blue.

Table 10. Hazard ratios and median rank with 95% CrI for ESKD in the subpopulation with diabetes and CKD, relative to placebo.

Class	Treatment	Hazard ratio Median (95% CrI)	Rank Median (95% CrI)
Placebo	Placebo	*Reference*	4 (3, 4)
SGLT-2i	Canagliflozin	0.68 (0.54, 0.86)	3 (2, 4)
	Dapagliflozin	0.48 (0.05, 4.37)	2 (1, 4)
DPP4-i	Saxagliptin	0.0006 (<0.0001, 0.64)	1 (1, 2)

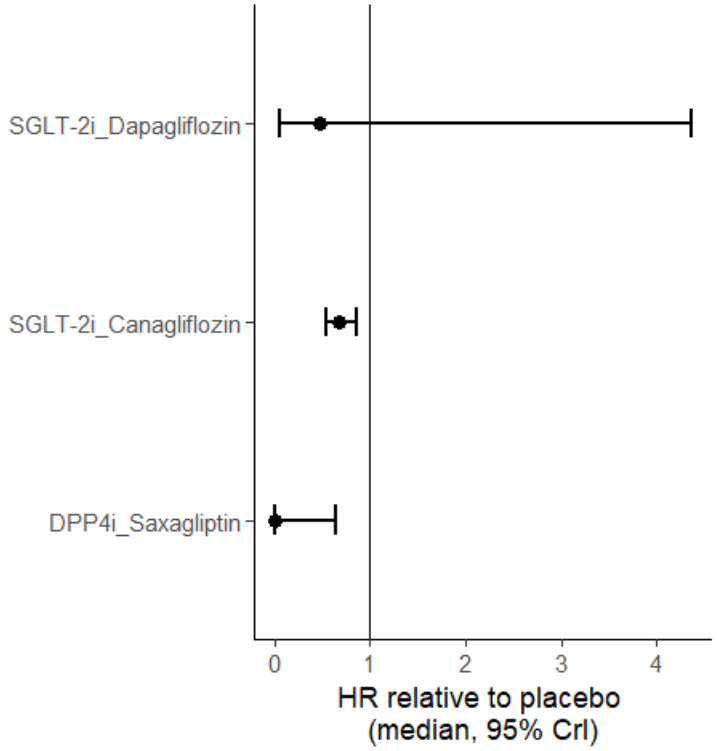

Figure 17. Hazard ratio for end-stage kidney disease, relative to placebo, in the population with T2D and CKD.

Definition of the subpopulation with heart failure (HF)

These analyses focus on people with type 2 diabetes mellitus (T2D) and heart failure (HF).

Studies in people with HF primarily were not included in the evidence, including where a T2D subgroup was reported.

Summary

In this subpopulation, there was RCT evidence for six outcomes:

- Change in HbA1c
- Hospitalisation for heart failure
- Cardiovascular mortality
- Three-point MACE
- Non-fatal stroke
- Non-fatal myocardial infarction

There was no evidence available for this subpopulation on treatment effects on weight change, unstable angina or end-stage kidney disease. Additionally, results for the impact of treatment on HbA1c are presented relative to insulin, because the network of evidence did not include placebo or metformin.

Fixed-effect models were preferred on model fit (total residual deviance and DIC) for all six outcomes (Subpopulation with diabetes and HF). All networks in this subpopulation were star networks, without loops providing indirect evidence. Therefore, inconsistency between direct and indirect evidence could not be assessed.

Subpopulation with HF: Change in HbA1c

In this dataset neither placebo nor metformin was represented in the network, therefore all comparisons are presented relative to insulin. The apparent loop in this network (Figure 18) was created from a three-armed trial, which will be consistent by design.

There was no clear evidence of differences in magnitude of HbA1c reduction in patients receiving sitagliptin and exenatide, relative to those receiving insulin. There was clear evidence of a smaller reduction in HbA1c in those study participants receiving liraglutide than in those participants receiving insulin. For the full table of active-active comparisons, see tab ‘Treatment Direct Effects’ in RQ1.2 results file HF HbA1c.

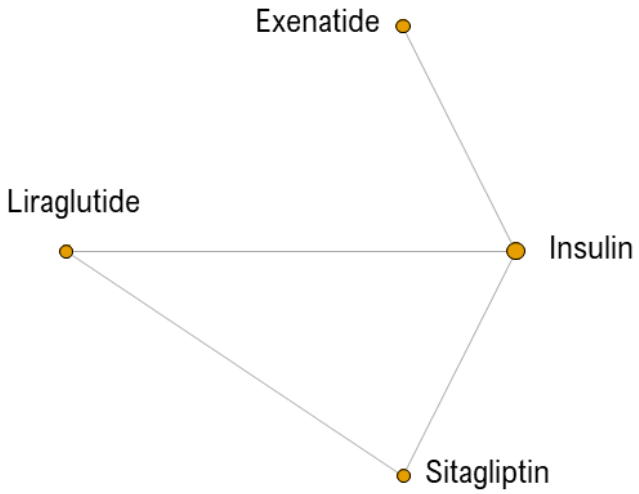

Figure 18. Network of evidence for change in HbA1c in the subpopulation of people with diabetes and heart failure.

*Table 11. Change in HbA1c relative to **insulin** and median rank with 95% CrI for the subpopulation with diabetes and HF, with fixed-effects structure on treatment effects.*

Class	Treatment	Change in HbA1c Median (95% CrI)	Rank Median (95% CrI)
Insulin	Insulin	*Reference*	2 (1, 3)
DPP4i	Sitagliptin	0.30 (-0.88, 1.48)	2 (1, 4)
GLP-1RA	Exenatide	0.20 (-0.99, 1.40)	2 (1, 3)
	Liraglutide	1.30 (0.10, 2.50)	4 (3, 4)

Subpopulation with HF: Cardiovascular (CV) mortality

This analysis included six trials of seven treatments (Figure 19). Five studies reported the hazard ratio, with three also reporting the number of events and one reporting the number of events alone. Between 4% and 10% of participants experienced CV mortality in those studies where the number of events was reported. Random-effect model structures were not considered for CV mortality because there was insufficient evidence to estimate the between-study variability.

For canagliflozin and alogliptin, there was evidence that participants receiving treatment saw a lower rate of CV mortality than those receiving placebo, though the 95% credible interval for the hazard ratio for both treatments included no effect (Table 12). There was considerable uncertainty around the conclusion of any protective effect of liraglutide, linagliptin and canagliflozin. The median of the estimated treatment effect for vildagliptin vs placebo indicates an increased hazard for those participants receiving vildagliptin; however, the credible interval is extremely wide and includes the probability of both harmful and protective effects. No clear differences were noted between active treatments (for the full table of active-active comparisons, see tab 'Treatment Direct Effects' in RQ1.2 results file HF CVM).

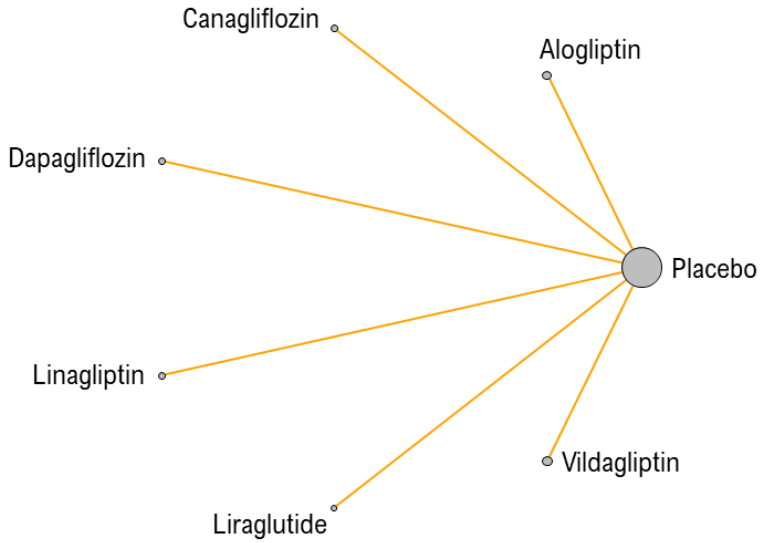

Figure 19. Network of evidence for cardiovascular mortality in the subpopulation of people with diabetes and heart failure.

Table 12. Hazard ratios and posterior median rank with 95% CrI for CV mortality in the subpopulation with diabetes and HF, relative to placebo.

Class	Treatment	Hazard ratio Median (95% CrI)	Rank Median (range)
Placebo	Placebo	*Reference*	5 (3, 7)
DPP4i	Alogliptin	0.77 (0.54, 1.09)	2 (1, 6)
	Linagliptin	0.96 (0.76, 1.26)	4 (2, 7)
	Vildagliptin	1.80 (0.54, 7.15)	7 (1, 7)
GLP-1RA	Liraglutide	0.85 (0.63, 1.15)	3 (1, 6)
SGLT-2i	Canagliflozin	0.72 (0.51, 1.02)	2 (1, 5)

	Dapagliflozin	1.01 (0.73, 1.39)	5 (2, 7)

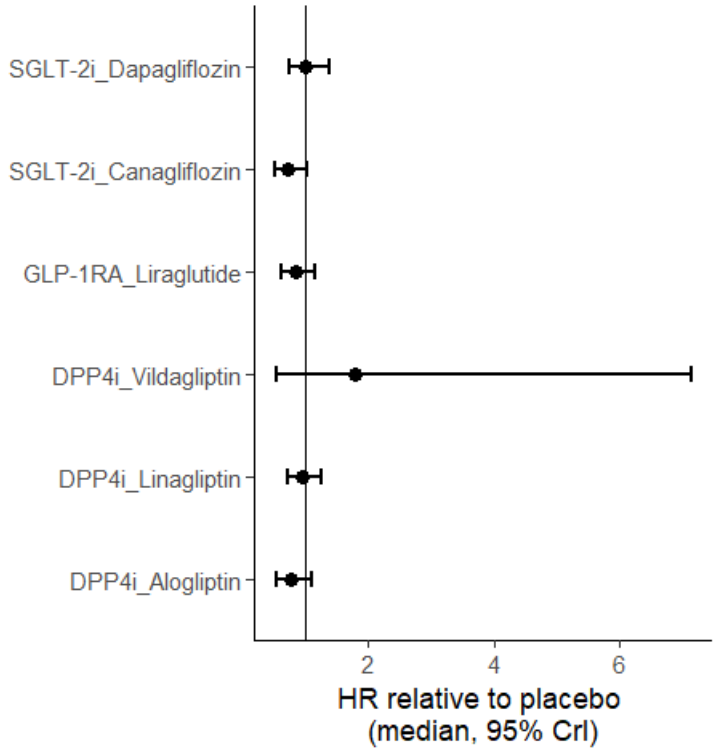

Figure 20. Hazard ratio for cardiovascular mortality, relative to placebo, in the population with T2D and HF.

Subpopulation with HF: hospitalisation for heart failure

This analysis included ten trials of ten treatments (Figure 21). Nine studies reported hazard ratios, with seven of these reporting both hazard ratios and the number of hospitalisations. Hospitalisation for heart failure affected between 6% and 13% of the participant populations of the studies where the number of events was reported.

There was clear evidence for a decreased hazard of hospitalisation for heart failure compared to placebo for all SGLT2-inhibitors in the network: canagliflozin, dapagliflozin and ertugliflozin (Table 13). The evidence was less clear for alogliptin, linagliptin, sitagliptin, vildagliptin, liraglutide and lixisenatide, with 95% credible intervals for these treatments including the probability of both protective and harmful effects when compared to placebo. There was clear evidence of a decreased hazard rate of hospitalisation for heart failure in those receiving canagliflozin, compared to those receiving sitagliptin, with a hazard ratio of 0.58 (95% CrI: 0.38, 0.89). Hazard rates were higher on liraglutide (median hazard ratio 1.60, 95% CrI: 1.06, 2.43) and alogliptin (1.63, 95% CrI: 1.02, 2.62) compared to canagliflozin and higher on sitagliptin compared to ertugliflozin (1.67, 95% CrI: 1.06, 2.62). For the full table of active-active comparisons, see tab ‘Treatment Direct Effects’ in RQ1.2 results file HF hosp.

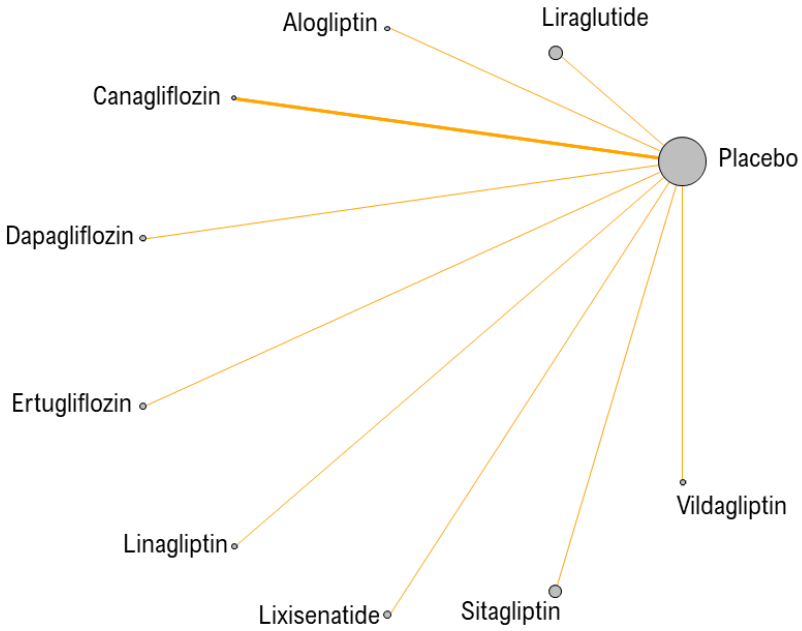

Figure 21. Network of evidence for hospitalisation with heart failure in the subpopulation of people with diabetes and heart failure.

Table 13. Hazard ratios and posterior median rank with 95% CrI for hospitalisation for heart failure in the subpopulation with diabetes and HF, relative to placebo.

Class	Treatment	Hazard ratio Median (95% CrI)	Rank Median (range)
DPP4i	Alogliptin	1.00 (0.71, 1.41)	7 (3, 10)
	Linagliptin	0.88 (0.68, 1.14)	5 (2, 9)
	Sitagliptin	1.05 (0.79, 1.39)	8 (4, 10)
	Vildagliptin	1.32 (0.58, 3.10)	10 (2, 10)
GLP-1RA	Liraglutide	0.98 (0.75, 1.28)	7 (3, 10)
	Lixisenatide	0.93 (0.66, 1.31)	6 (2, 10)
SGLT-2i	Canagliflozin	0.61 (0.45, 0.84)	2 (1, 4)
	Dapagliflozin	0.73 (0.55, 0.96)	3 (1, 6)
	Ertugliflozin	0.63 (0.44, 0.90)	2 (1, 5)

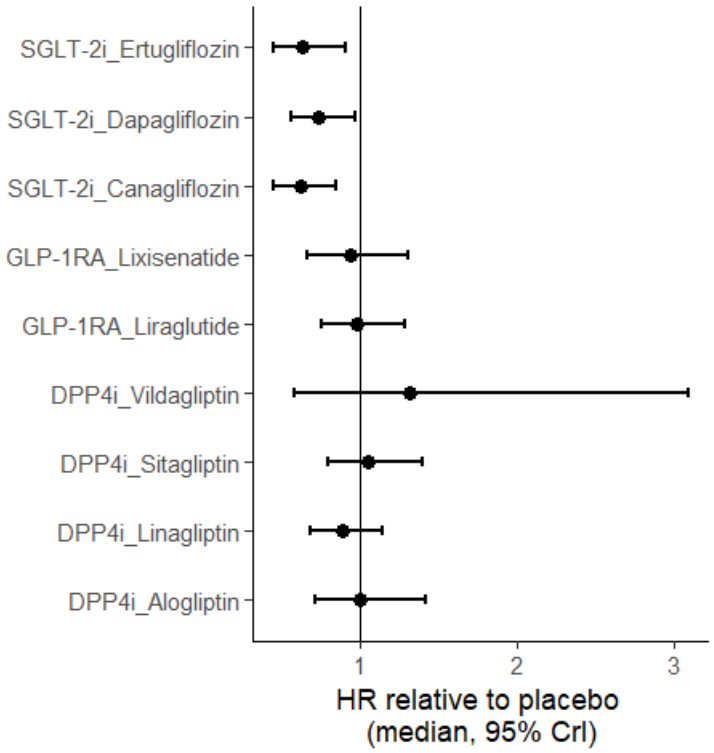

Figure 22. Hazard ratio for hospitalisation for H,F relative to placebo, in the population with T2D and HF.

Subpopulation with HF: three-point MACE

This analysis included eight trials of eight treatments (Figure 23). All studies reported hazard ratios, which were preferred over event counts for synthesis in the NMA. Six studies also reported the number of three-point MACE events, experiencing one of the events in this composite outcome was relatively common in this population, affecting between 12% and 19% of study participants.

There was weak evidence for a decreased hazard of three-point MACE compared to placebo for liraglutide and canagliflozin (Table 14). The evidence was less clear for alogliptin, exenatide, subcutaneous semaglutide, dapagliflozin and ertugliflozin, with the median treatment effect indicating no difference between the active treatment and placebo, and 95% credible intervals for these treatments including the probability of both protective and harmful effects. No clear differences were noted between active treatments (for the full table of active-active comparisons, see tab 'Treatment Direct Effects' in RQ1.2 results file HF MACE3).

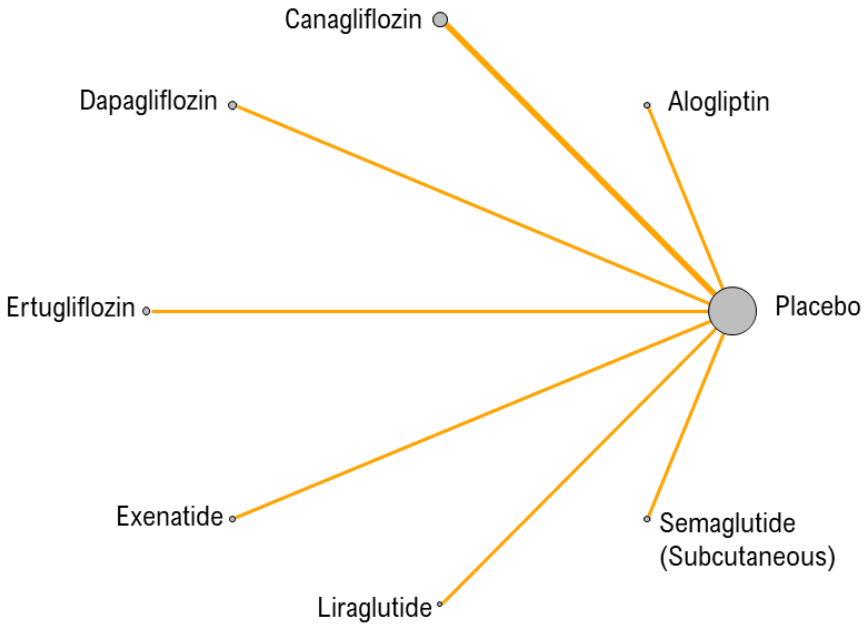

Figure 23. Network of evidence for three-point MACE in the subpopulation of people with diabetes and heart failure.

Table 14. Hazard ratios and median rank with 95% CrI for three-point MACE in the subpopulation with diabetes and HF, relative to placebo.

Class	Treatment	Hazard ratio Median (95% CrI)	Rank Median (range)
DPP4i	Alogliptin	0.94 (0.74, 1.20)	7 (2, 8)
GLP-1RA	Exenatide	0.97 (0.81, 1.16)	5 (2, 8)
	Liraglutide	0.81 (0.65, 1.02)	2 (1, 6)
	Semaglutide (subcutaneous)	1.03 (0.64, 1.66)	2 (1, 6)
SGLT-2i	Canagliflozin	0.84 (0.67, 1.04)	6 (1, 8)
	Dapagliflozin	1.01 (0.81, 1.26)	4 (1, 8)
	Ertugliflozin	1.05 (0.82, 1.35)	6 (2, 8)

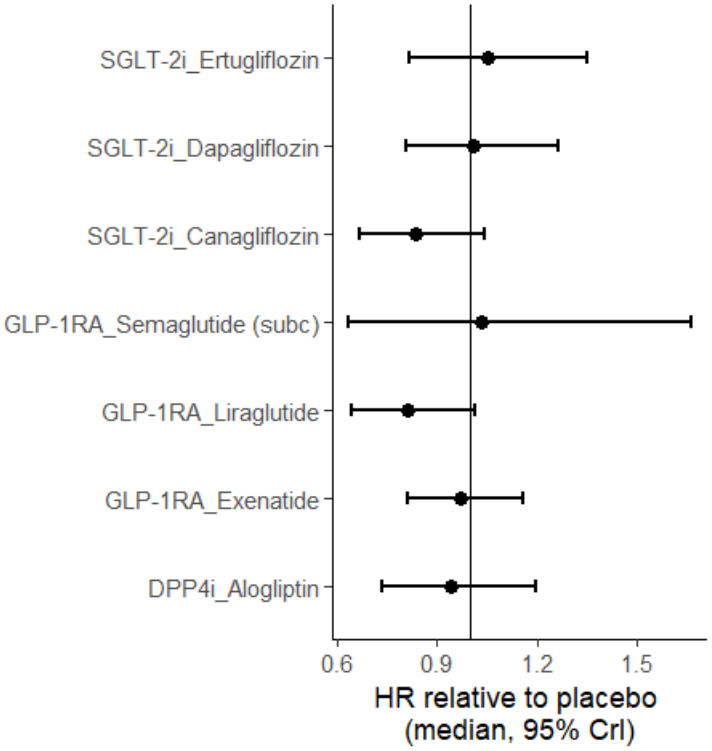

Figure 24. Hazard ratio for three-point MACE relative to placebo in the population with T2D and HF.

Subpopulation with HF: non-fatal stroke

This analysis included four trials of five treatments (Figure 25). All studies reported the number of events, with two also reporting hazard ratios. Between 1% and 4% of study participants experienced non-fatal stroke. Random-effect model structures were not considered for non-fatal stroke because there was insufficient evidence to estimate the between-study variability.

There was weak evidence for a decreased hazard of non-fatal stroke compared to placebo for vildagliptin (Table 15). This was based upon evidence from McMurray 2018, which observed four non-fatal strokes in the 126 people on the placebo arm and one event in 128 people on the vildagliptin arm. The evidence for alogliptin, liraglutide and dapagliflozin was more uncertain, with 95% credible intervals for these treatments including the probability of both protective and harmful effects. No precisely estimated differences were noted between active treatments (for the full table of active-active comparisons, see tab 'Treatment Direct Effects' in RQ1.2 results file HF stroke).

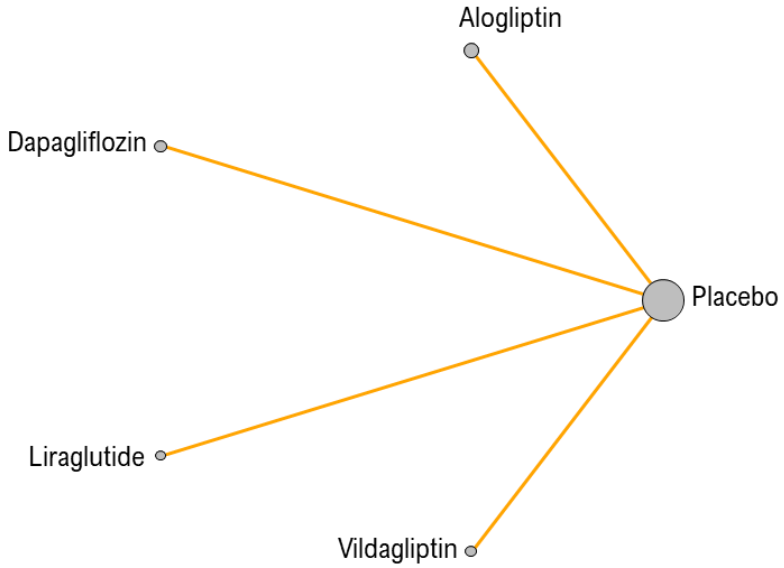

Figure 25. Network of evidence for non-fatal stroke in the subpopulation of people with diabetes and heart failure.

Table 15. Hazard ratios and median rank with 95% CrI for non-fatal stroke in the subpopulation with diabetes and HF, relative to placebo.

Class	Treatment	Hazard ratio Median (95% CrI)	Rank Median (range)
DPP4i	Alogliptin	1.85 (0.70, 5.56)	5 (2, 5)
	Vildagliptin	0.04 (<0.01, 1.12)	1 (1, 3)
GLP-1RA	Liraglutide	0.89 (0.53, 1.50)	2 (1, 5)
SGLT-2i	Dapagliflozin	1.21 (0.76, 1.90)	4 (2, 5)

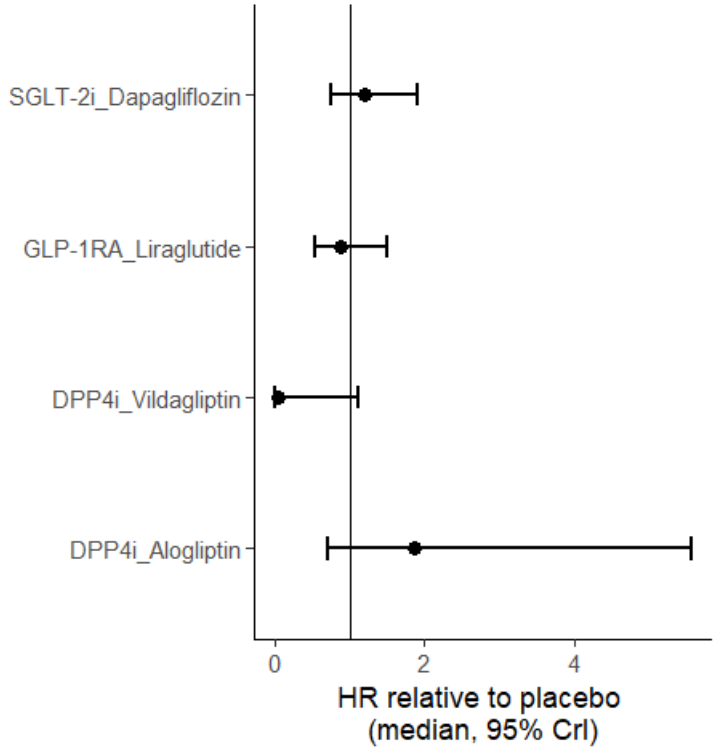

Figure 26. Hazard ratio for non-fatal stroke, relative to placebo, in the population with T2D and HF.

Subpopulation with HF: non-fatal myocardial infarction (MI)

This analysis included three trials of four treatments (Figure 27). All studies reported the number of events, with two also reporting hazard ratios. Between 7% and 9% of study participants experienced non-fatal MI. Random-effect model structures were not considered for non-fatal MI because there was insufficient evidence to estimate the between-study variability.

There was weak evidence for a decreased hazard of non-fatal MI compared to placebo for liraglutide (Table 16). The evidence for alogliptin and dapagliflozin was more uncertain, with 95% credible intervals for these treatments including the probability of both protective and harmful effects. No clear differences were noted between active treatments (for the full table of active-active comparisons, see tab 'Treatment Direct Effects' in RQ1.2 results file HF MI).

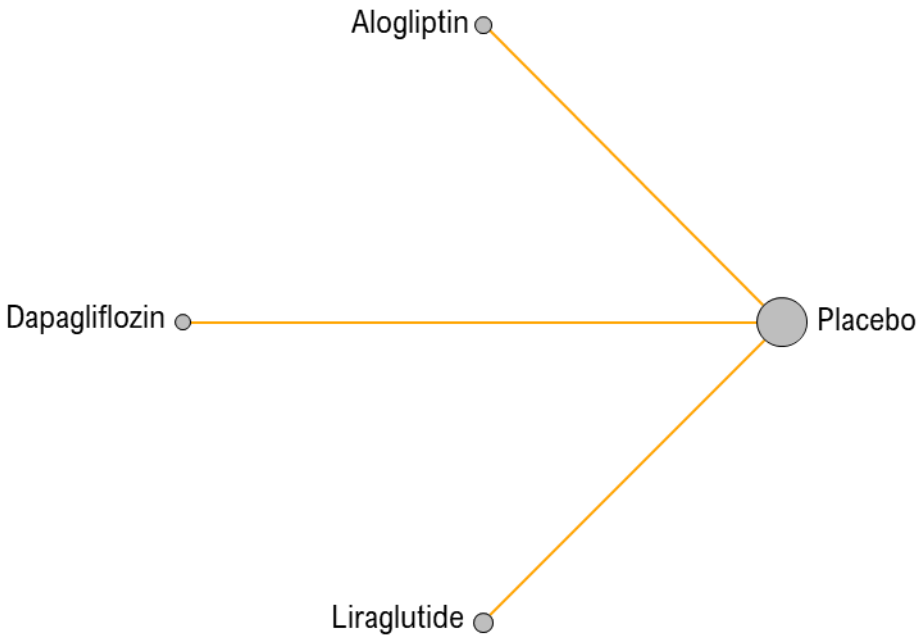

Figure 27. Network of evidence for non-fatal MI in the subpopulation of people with diabetes and heart failure.

Table 16. Hazard ratios and median rank with 95% CrI for non-fatal MI in the subpopulation with diabetes and HF, relative to placebo.

Class	Treatment	Hazard ratio Median (95% CrI)	Rank Median (range)
DPP4i	Alogliptin	1.04 (0.74, 1.46)	4 (1, 4)
GLP-1RA	Liraglutide	0.74 (0.52, 1.06)	1 (1, 4)
SGLT-2i	Dapagliflozin	0.85 (0.61, 1.18)	2 (1, 4)

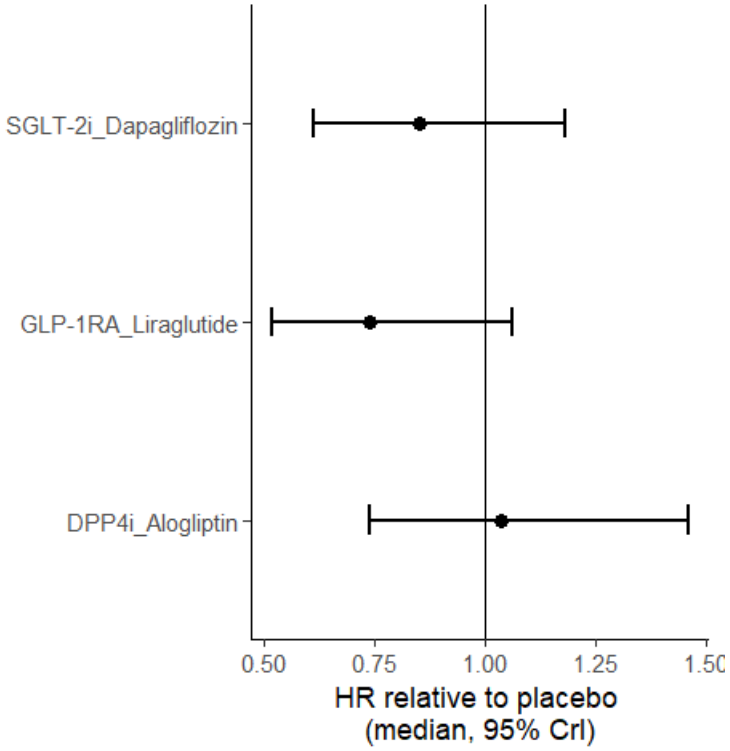

Figure 28. Hazard ratio for non-fatal MI relative to placebo in the population with T2D and HF.

Definition of the subpopulation with atherosclerotic (AS) CVD

These analyses focus on people with type 2 diabetes mellitus (T2D) and atherosclerotic cardiovascular disease (AS-CVD)

Studies in people with AS-CVD primarily were not included in the evidence, including where a T2D subgroup was reported.

Summary

In this population, there was RCT evidence for eight outcomes:

- Change in HbA1c
- Change in body weight
- Cardiovascular mortality
- Hospitalisation for heart failure
- Three-point MACE
- Non-fatal stroke
- Non-fatal myocardial infarction
- Unstable angina

Fixed-effect models were preferred on model fit (total residual deviance and DIC) for all outcomes except change in HbA1c. Most networks in this subpopulation were star networks, with RCT evidence trialling active treatments against placebo and without loops providing indirect evidence. Therefore, inconsistency between direct and indirect evidence could not be assessed for all outcomes except change in HbA1c.

Subpopulation with AS-CVD: Change in HbA1c

This analysis included 18 studies of 14 treatments (Figure 29). The random-effects model, in which study-specific relative effects are assumed to be exchangeable, was preferred on improved model fit (lower residual deviance and DIC, Table 39).

This is due to the heterogeneity in study-specific relative effects present in the dataset. Between-study SD was 0.31 (95% CrI: 0.16, 0.69), representing moderate heterogeneity in units of % HbA1c.

There is evidence that treatment with pioglitazone, glimepiride, tirzepatide and dapagliflozin results in greater reductions in % HbA1c than treatment with placebo (Figure 30). The evidence for a reduction in % HbA1c relative to placebo was weaker for insulin; the SGLT2-inhibitors ertugliflozin and empagliflozin; the DPP4-inhibitors alogliptin, sitagliptin and vildagliptin; and the GLP1-receptor agonists exenatide, liraglutide and lixisenatide. For these treatments there was greater uncertainty around the effect of treatment. For alogliptin and sitagliptin, the median treatment effect was estimated to be closer to zero (no difference relative to placebo), whilst for liraglutide and insulin the majority of the posterior distribution was on the left of the line of no effect, indicating a reduction in % HbA1c relative to placebo.

Estimates of model fit and between-study SD were similar between NMA and UME models, suggesting that there was little inconsistency in this network (Table 25). No study arms were

identified as containing direct evidence in conflict with indirect evidence from the network (Figure 31).

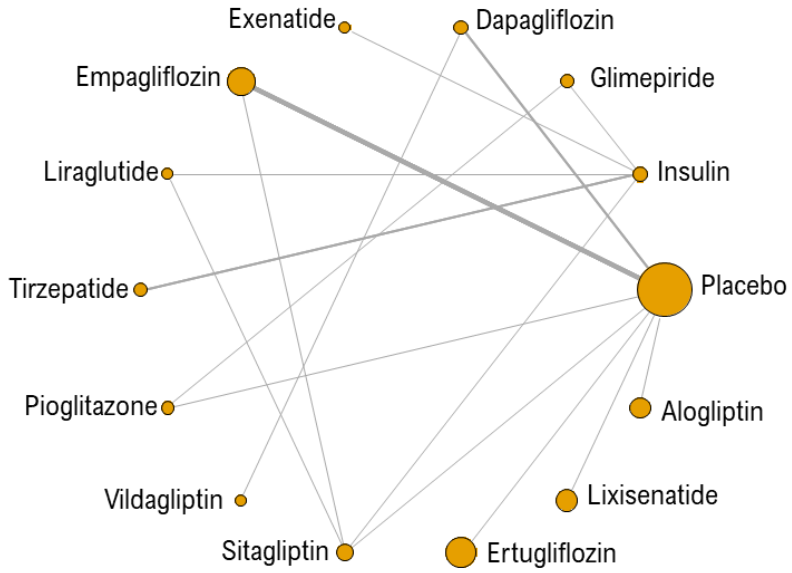

Figure 29. Network of evidence for change in HbA1c in the subpopulation with diabetes and AS-CVD.

Table 17. Change in HbA1c relative to placebo for 13 active treatments in the network, with random-effects structure on treatment effects (for the subpopulation with AS-CVD).

Class	Treatment	Mean difference, Change in HbA1c
DPP4i	Alogliptin	-0.36 (-1.12, 0.38)
	Sitagliptin	-0.23 (-0.75, 0.34)
	Vildagliptin	-0.79 (-1.96, 0.37)
GIPGLP	Tirzepatide	-1.83 (-3.12, -0.61)
GLP-1RA	Exenatide	-0.57 (-2.39, 1.21)
	Liraglutide	-1.11 (-2.29, 0.10)
	Lixisenatide	-0.27 (-1.01, 0.48)
Insulin	Insulin	-0.86 (-1.98, 0.22)
SGLT-2i	Dapagliflozin	-0.58 (-1.11, -0.05)
	Empagliflozin	-0.24 (-0.63, 0.16)
	Ertugliflozin	-0.17 (-0.92, 0.58)
Sulfonylurea	Glimepiride	-1.14 (-2.20, -0.13)
TZD	Pioglitazone	-1.17 (-2.08, -0.27)

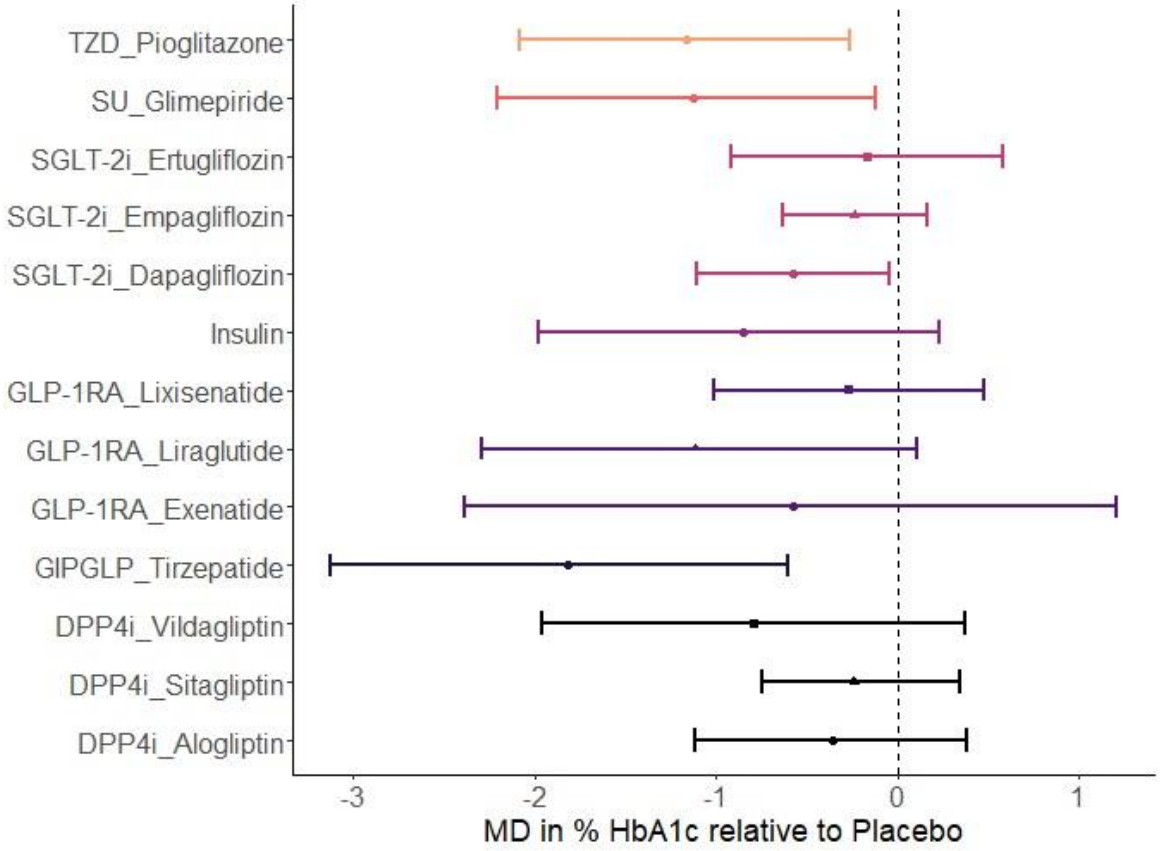

Figure 30. Mean difference (MD) in % HbA1c relative to placebo in the subpopulation with diabetes and AS-CVD. Mean treatment effects are shown as points, with the 95% credible interval. Line colour corresponds to treatment class, with point shape distinguishing different treatments within the same class.

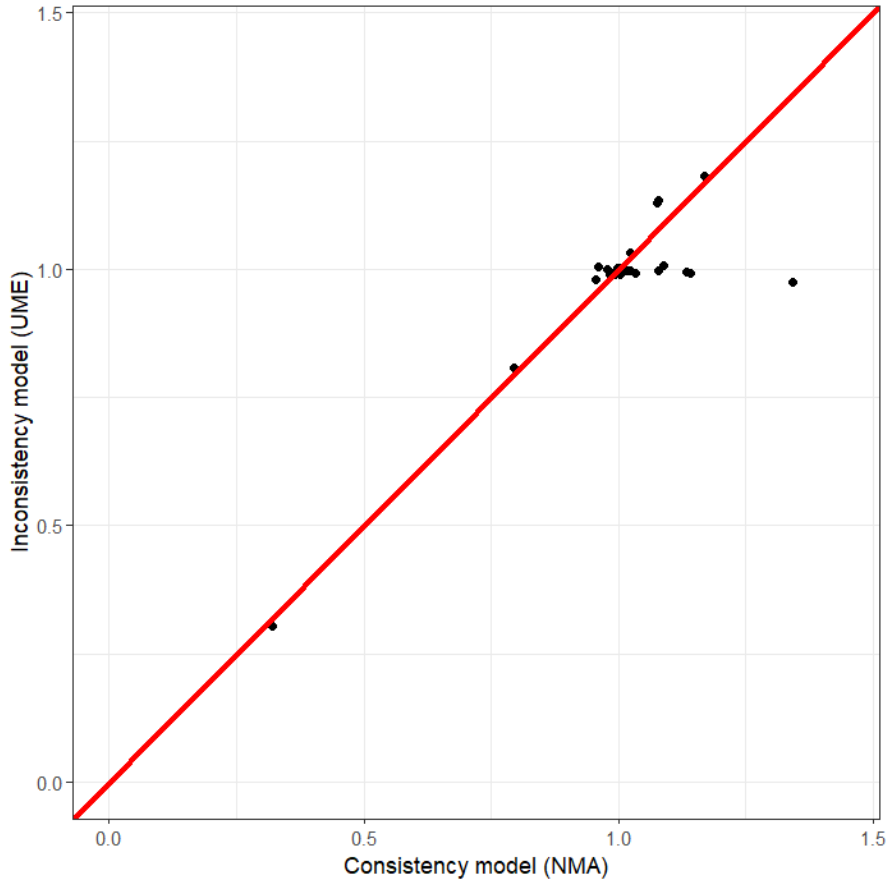

Figure 31. Dev-dev plot showing residual deviance contribution for each study arm under the NMA and UME models. Study arms that fit poorly under the consistency model would show relatively high deviance, and appear as points in the lower right-hand region.

Subpopulation with AS-CVD: Change in weight

This analysis included five trials of five treatments (Figure 32). Fixed- and random-effect models were run, however there was only one edge informed by more than a single study (dapagliflozin vs placebo) and there was very little evidence from which to estimate the between-study heterogeneity, leading to poor convergence in the RE model (Table 40).

There was clear evidence of a very small reduction in weight for those patients on dapagliflozin (Table 18). The evidence for empagliflozin and sitagliptin was more uncertain, with the 95% credibility interval supporting the probability of weight-loss, weight-neutral and weight-gain effects. The evidence for vildagliptin was consistent with a very small weight gain on treatment, relative to the weight change on placebo. There was evidence of weight gain on vildagliptin relative to dapagliflozin, with a mean difference of 4.4kg (95% CrI 2.63kg, 6.19kg).

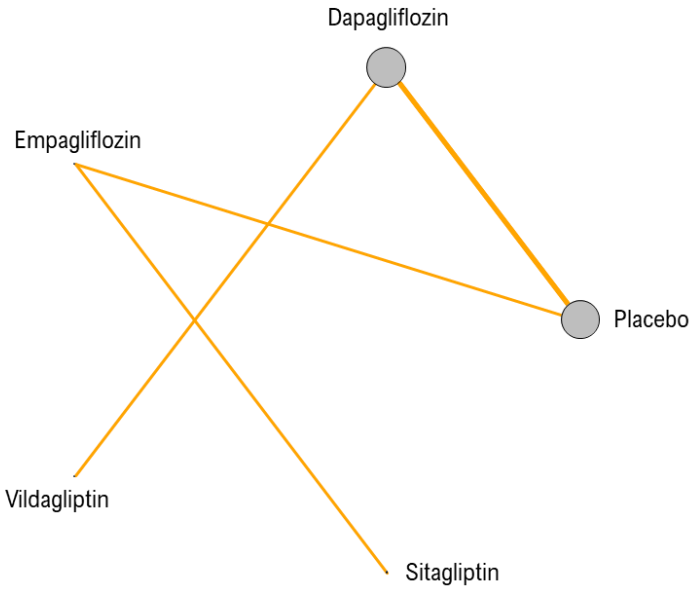

Figure 32. Network of evidence for proportional weight change in the subpopulation of people with diabetes and AS-CVD.

Table 18. Predicted weights at follow-up and treatment effects as ratios and mean difference relative to placebo for the four active treatments in the network for the subpopulation of people with T2D and AS-CVD. This was a proportional model in which the mean difference will change depending on weight. Therefore results are given for patients at two representative weights on placebo: 60kg and 90kg.

Class	Treatment	Assuming a final weight on placebo of		Ratio	Mean difference [1] (kg) Mean, 95% CrI
		60kg	90kg		

[1] Mean difference in kg between CFB on treatment arm and CFB on placebo arm, given baseline weight of 93.2kg and mean CFB on placebo of -0.6kg, as reported by Leiter 2014, the study with the greatest number of participants on the placebo arm.

SGLT-2i	Dapagliflozin	58.6 (58.3, 58.8)	87.8 (87.5, 88.2)	0.98 (0.97, 0.98)	-2.23 (-2.58, -1.88)
	Empagliflozin	61.1 (55.7, 66.7)	91.6 (83.6, 100.0)	1.02 (0.93, 1.11)	1.60 (-6.65, 10.33)
DPP4i	Sitagliptin	63.3 (56.8, 70.4)	95.0 (85.2, 105.6)	1.05 (0.95, 1.17)	4.89 (-4.95, 16.10)
	Vildagliptin	61.4 (60.2, 62.6)	92.1 (90.4, 93.9)	1.02 (1.00, 1.04)	2.17 (0.37, 4.01)

Subpopulation with AS-CVD: Cardiovascular (CV) mortality

This analysis included twelve trials of ten treatments (Figure 33). Eight studies reported hazard ratios, with four reporting the number of events alone. CV mortality was rare in this population, with between 0.6% and 6% of participants affected in studies where it was reported.

There was clear evidence for a decreased hazard of CV mortality compared to placebo for empagliflozin (Table 19). The evidence for canagliflozin and alogliptin was more uncertain, but suggested a decreased hazard of CV mortality compared to placebo. The evidence for sitagliptin, glimepiride, lixisenatide, dapagliflozin, empagliflozin, ertugliflozin and pioglitazone is consistent with there being no difference compared with placebo in hazard of CV mortality in this population. The 95% credible intervals for these treatments included the probability of both protective and harmful effects. Hazard rates were higher on ertugliflozin (median hazard ratio 1.49, 95% CrI: 1.12, 1.99), dapagliflozin (1.54, 95% CrI: 1.13, 2.10), lixisenatide (1.58, 95% CrI: 1.15, 2.17), sitagliptin (1.66, 95% CrI: 1.27, 2.17), canagliflozin (1.39, 95% CrI: 1.02, 1.89) and pioglitazone (1.52, 95% CrI: 1.09, 2.11) than on empagliflozin. For the full table of active-active comparisons, please see tab 'Treatment Direct Effects' in RQ1.2 results file CVD CVM.

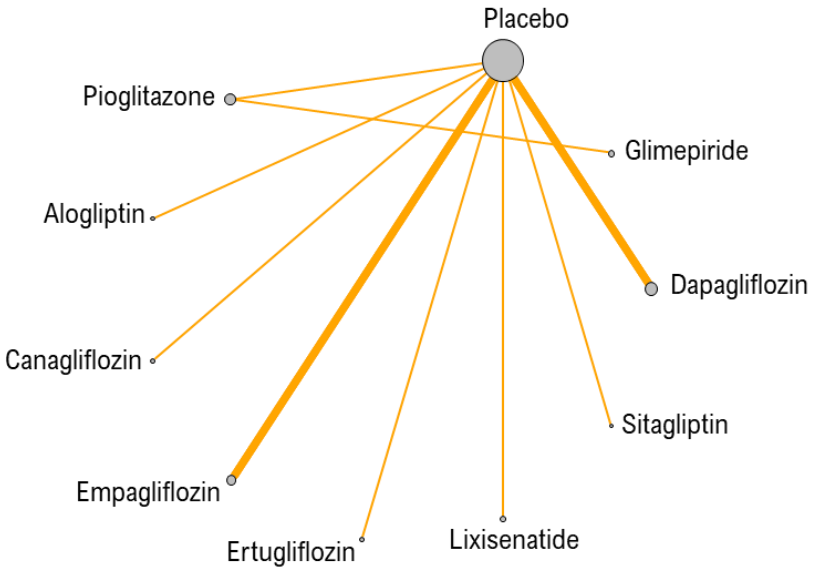

Figure 33. Network of evidence for cardiovascular mortality in the subpopulation of people with diabetes and AS-CVD.

Table 19. Hazard ratios and median rank with 95% CrI for CV mortality in the subpopulation with diabetes and AS-CVD, relative to placebo.

Class	Treatment	Hazard ratio Median (95% CrI)	Rank Median (range)
DPP4i	Alogliptin	0.85 (0.66, 1.10)	4 (2, 10)
	Sitagliptin	1.03 (0.89, 1.19)	9 (4, 10)
Sulfonylurea	Glimepiride	0.28 (0.01, 2.36)	1 (1, 10)
GLP-1RA	Lixisenatide	0.98 (0.78, 1.23)	7 (3, 10)
SGLT-2i	Canagliflozin	0.86 (0.70, 1.06)	4 (2, 9)
	Dapagliflozin	0.95 (0.77, 1.19)	6 (3, 10)
	Empagliflozin	0.62 (0.49, 0.78)	2 (1, 3)
	Ertugliflozin	0.92 (0.77, 1.11)	6 (3, 10)
TZD	Pioglitazone	0.94 (0.74, 1.20)	6 (3, 10)

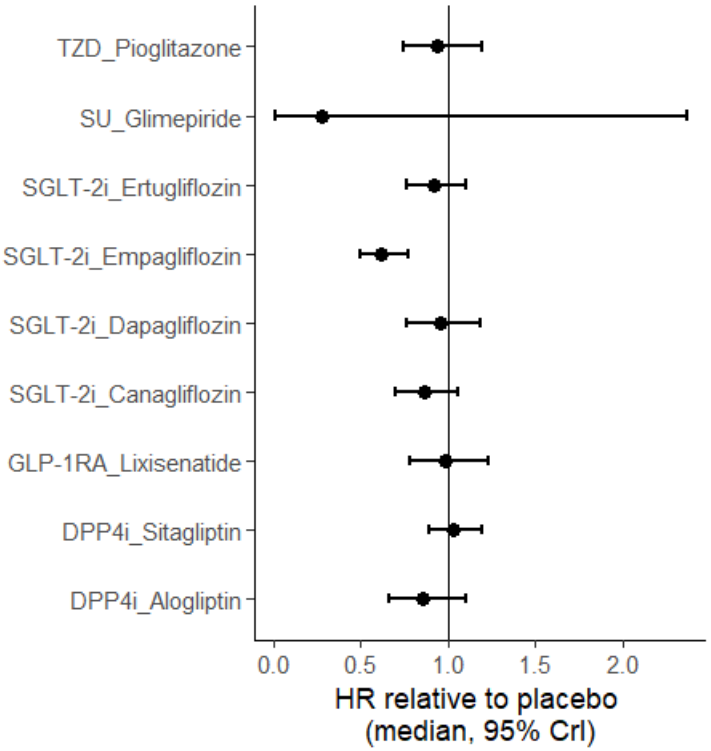

Figure 34. Hazard ratio for cardiovascular mortality, relative to placebo, in the population with T2D and AS-CVD.

Subpopulation with AS-CVD: hospitalisation for heart failure

This analysis included eleven trials of ten treatments (Figure 35). Seven trials reported hazard ratios, four reported number of events alone. Hospitalisation for heart failure was a rare outcome in this population, with less than 5% of participants experiencing hospitalisation in all studies where the number of events was reported.

There was clear evidence for decreased hazard of hospitalisation for heart failure compared with placebo for all SGLT2-inhibitors in the network: canagliflozin, dapagliflozin, empagliflozin and

ertugliflozin (Table 20). There was clear evidence for an increased hazard of hospitalisation for heart failure compared with placebo for pioglitazone. The evidence for alogliptin, sitagliptin, glimepiride and lixisenatide was more uncertain with the 95% credible intervals for these treatments including the probability of both protective and harmful effects. The estimate for glimepiride was particularly uncertain, indicated in the range of rank including all possible rank positions. This was estimated from a single study, Nissen 2008, which trialled pioglitazone against glimepiride. This study reported five events in 273 people on the glimepiride arm and four events in 270 people on the pioglitazone arm.

There was clear evidence for a decreased hazard with canagliflozin, compared with both alogliptin (median hazard ratio 0.57, 95% CrI: 0.38, 0.85) and sitagliptin (0.68, 95% CrI: 0.48, 0.95). Hazard rates were higher on alogliptin than on dapagliflozin (1.5, 95% CrI: 1.06, 2.14), empagliflozin (1.83, 95% CrI: 1.24, 2.69) and ertugliflozin (1.7, 95% CrI: 1.16, 2.48); higher on sitagliptin than on empagliflozin (1.54, 95% CrI: 1.11, 2.12) and ertugliflozin (1.43, 95% CrI: 1.04, 1.96); and higher on lixisenatide than empagliflozin (1.48, 95% CrI: 1.03, 2.13). Hazard rates were higher on pioglitazone than on canagliflozin (2.08, 95% CrI: 1.43, 3.03), dapagliflozin (1.78, 95% CrI: 1.29, 2.47), empagliflozin (2.17, 95% CrI: 1.51, 3.12), ertugliflozin (2.01, 95% CrI: 1.41, 2.87) and sitagliptin (1.41, 95% CrI: 1.04, 1.92). For the full table of active-active comparisons, see tab 'Treatment Direct Effects' in RQ1.2 results file CVD hosp.

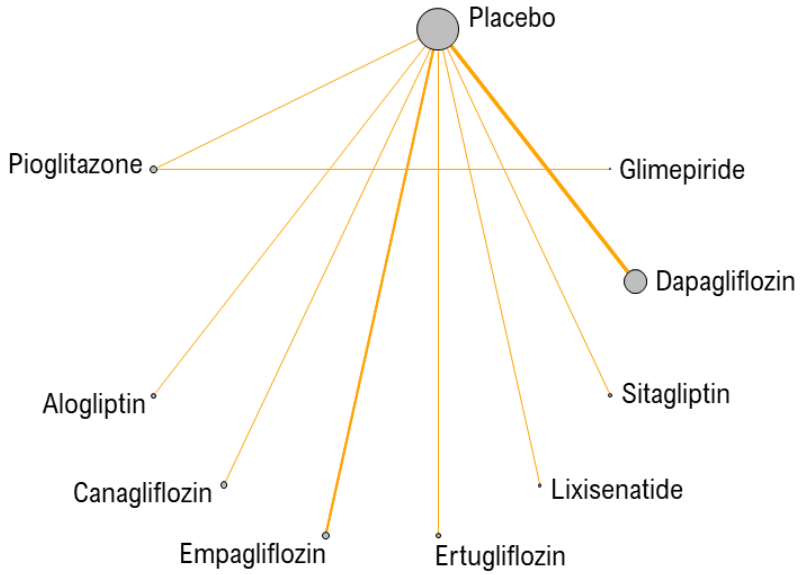

Figure 35. Network of evidence for hospitalisation for heart failure in the subpopulation of people with diabetes and AS-CVD.

Table 20. Hazard ratios and median rank with 95% CrI for hospitalisation for heart failure in the subpopulation with diabetes and AS-CVD, relative to placebo.

Class	Treatment	Hazard ratio Median (95% CrI)	Rank Median (range)
DPP4i	Alogliptin	1.19 (0.90, 1.58)	8 (5, 10)
	Sitagliptin	1.00 (0.83, 1.20)	6 (4, 9)
Sulfonylurea	Glimepiride	1.79 (0.45, 7.42)	10 (1, 10)

GLP-1RA	Lixisenatide	0.96 (0.75, 1.23)	6 (3, 9)
SGLT-2i	Canagliflozin	0.68 (0.51, 0.90)	2 (1, 5)
	Dapagliflozin	0.79 (0.64, 0.98)	4 (1, 6)
	Empagliflozin	0.65 (0.50, 0.85)	2 (1, 4)
	Ertugliflozin	0.70 (0.54, 0.91)	3 (1, 5)
TZD	Pioglitazone	1.41 (1.10, 1.80)	9 (8, 10)

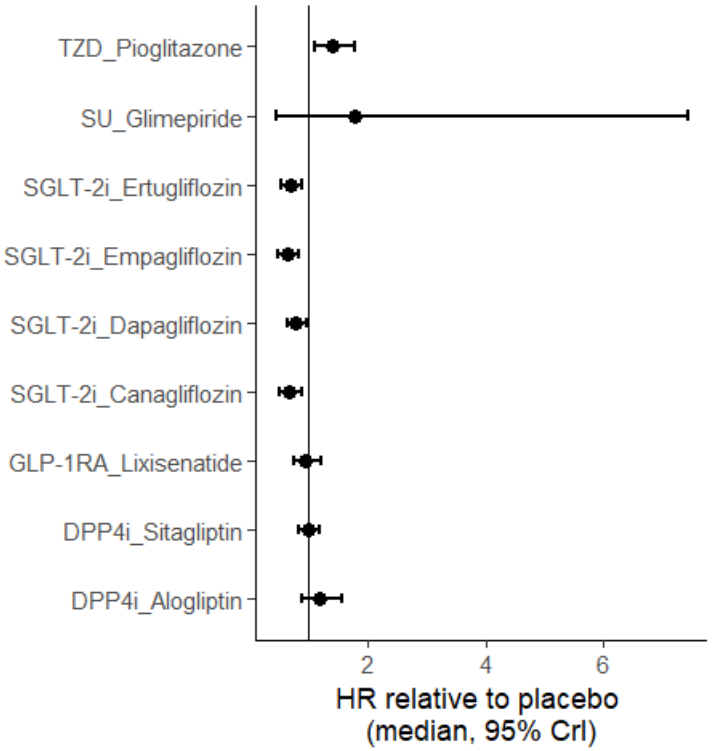

Figure 36. Hazard ratio for hospitalisation for heart failure, relative to placebo, in the population with T2D and AS-CVD.

Subpopulation with AS-CVD: three-point MACE

This analysis included 12 trials of 13 treatments (Figure 37). Nine studies reported hazard ratios, of which eight also reported the number of events, a further three studies reported number of events alone. In Nissen 2008, which trialled pioglitazone against glimepiride, these events were rare with only 2% of participants experiencing an event. In the rest of the dataset, these events were relatively common: between 8% and 19% of study participants experienced an event included in the composite three-point MACE. Random-effect model structures were not considered for three-point MACE, because there was insufficient evidence to estimate the between-study variability.

There was clear evidence suggesting lower hazard of three-point MACE events vs placebo for canagliflozin, exenatide and pioglitazone (Table 21). There was weaker evidence for a lower hazard of three-point MACE vs placebo for dulaglutide, dapagliflozin and empagliflozin. There was uncertain evidence for a very small reduction in hazard of three-point MACE vs placebo for alogliptin, saxagliptin, sitagliptin and ertugliflozin. Many treatments were similarly effective, leading to a high degree of overlapping within rankings. There was clear evidence of a decreased hazard rate in those receiving canagliflozin, compared to those receiving sitagliptin, with a hazard ratio of 0.83 (95% CrI:

0.70, 0.99) (for the full table of active-active comparisons, see tab 'Treatment Direct Effects' in RQ1.2 results file CVD MACE3).

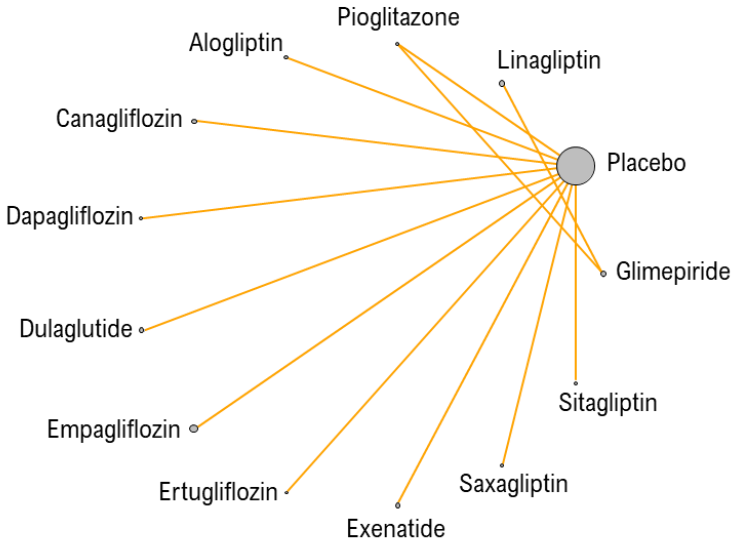

Figure 37. Network of evidence for three-point MACE in the subpopulation of people with diabetes and AS-CVD.

Table 21. Hazard ratios and posterior median rank with 95% CrI for three-point MACE in the subpopulation with diabetes and AS-CVD, relative to placebo.

Class	Treatment	Hazard ratio Median (95% CrI)	Rank Median (range)
DPP4i	Alogliptin	0.95 (0.81, 1.12)	9 (3, 13)
	Linagliptin	0.88 (0.28, 3.22)	5 (1, 13)
	Saxagliptin	0.96 (0.85, 1.08)	9 (4, 13)
	Sitagliptin	0.99 (0.89, 1.10)	10 (5, 13)
Sulfonylurea	Glimepiride	0.94 (0.30, 3.36)	8 (1, 13)
GLP-1RA	Dulaglutide	0.87 (0.74, 1.02)	5 (1, 11)
	Exenatide	0.90 (0.82, 0.99)	6 (2, 11)
SGLT-2i	Canagliflozin	0.82 (0.71, 0.94)	3 (1, 8)
	Dapagliflozin	0.90 (0.79, 1.02)	6 (2, 12)
	Empagliflozin	0.86 (0.74, 1.00)	4 (1, 11)
	Ertugliflozin	0.97 (0.85, 1.11)	9 (4, 13)
TZD	Pioglitazone	0.82 (0.70, 0.97)	3 (1, 9)

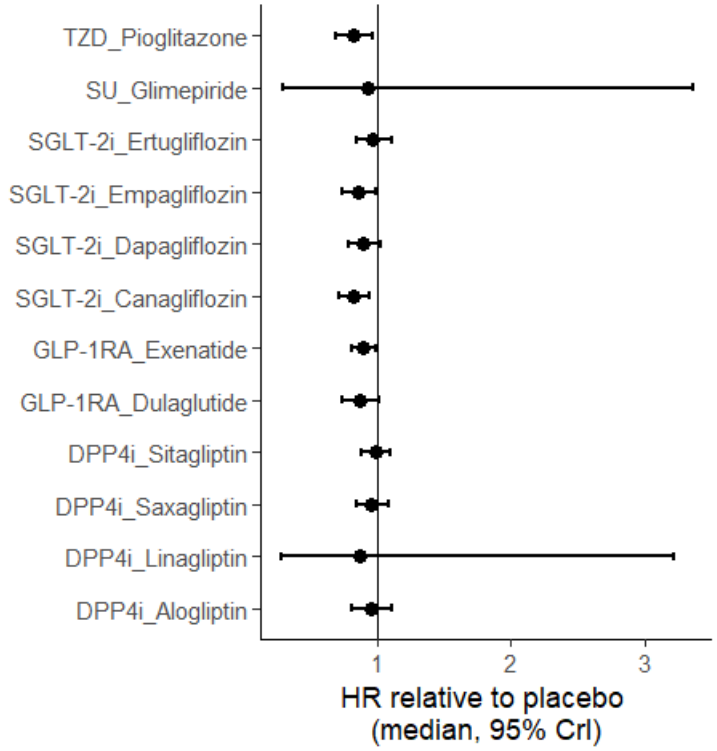

Figure 38. Hazard ratio for three-point MACE, relative to placebo, in the population with T2D and AS-CVD.

Subpopulation with AS-CVD: non-fatal stroke

This analysis included eight trials of eight treatments (Figure 39). Six studies reported hazard ratios, with a further two studies reporting the number of events. In Nissen 2008, which trialled pioglitazone against glimepiride, and Leiter 2014, which trialled dapagliflozin against placebo, non-fatal stroke was very rare: rates were equivalent to 2 events per 1000 participants. In the rest of the dataset, non-fatal stroke was rare: between 1% and 4% of study participants experienced non-fatal stroke. Random-effect model structures were not considered for non-fatal stroke, because there was insufficient evidence to estimate the between-study variability.

The evidence for alogliptin, dapagliflozin, empagliflozin and ertugliflozin was uncertain, with the 95% credible intervals for these treatments including the probability of both protective and harmful effects relative to placebo (Table 22). The estimate for glimepiride was not estimable and was based upon a single event on the glimepiride arm of Nissen 2008 (273 participants) compared with zero events on the pioglitazone arm (270 participants). There was weak evidence for a lower hazard of non-fatal stroke in those receiving pioglitazone, relative to those receiving placebo. There was clear evidence of a decreased hazard rate in those receiving pioglitazone, compared to those receiving empagliflozin, with a hazard ratio of 0.65 (95% CrI 0.43, 0.98) (for the full table of active-active comparisons, see tab ‘Treatment Direct Effects’ in RQ1.2 results file CVD stroke).

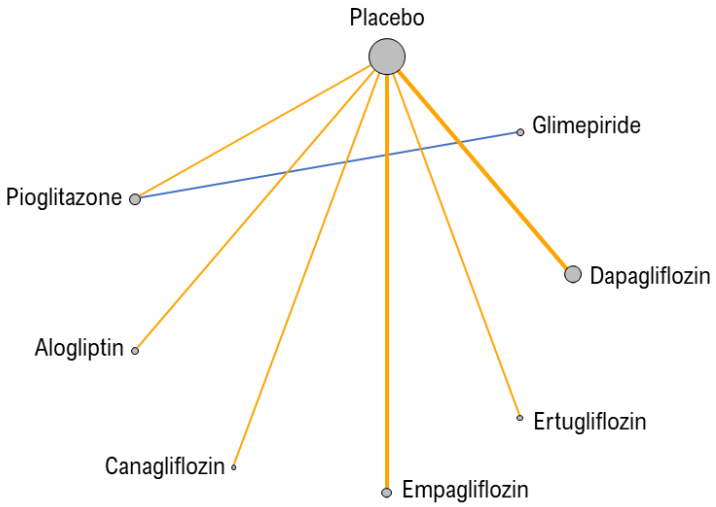

Figure 39. Network of evidence for non-fatal stroke, in the subpopulation of people with diabetes and AS-CVD. Network edges where data were sparse (zero events on one study arm) are shown in blue.

Table 22. Hazard ratios for non-fatal stroke in the subpopulation with diabetes and AS-CVD, relative to placebo. Median and range of rankings are also presented - these are less easy to interpret where hazard ratios are uncertain.

Class	Treatment	Hazard ratio Median (95% CrI)	Rank Median (range)
DPP4i	Alogliptin	0.92 (0.56, 1.51)	3 (1, 7)
Sulfonylurea	Glimepiride	*Not estimable*	*Not estimable*
SGLT-2i	Canagliflozin	0.88 (0.67, 1.16)	3 (1, 7)
	Dapagliflozin	0.97 (0.76, 1.22)	4 (1, 7)
	Empagliflozin	1.24 (0.92, 1.65)	7 (3, 8)
	Ertugliflozin	1.00 (0.76, 1.32)	5 (1, 7)
TZD	Pioglitazone	0.81 (0.61, 1.07)	2 (1, 6)

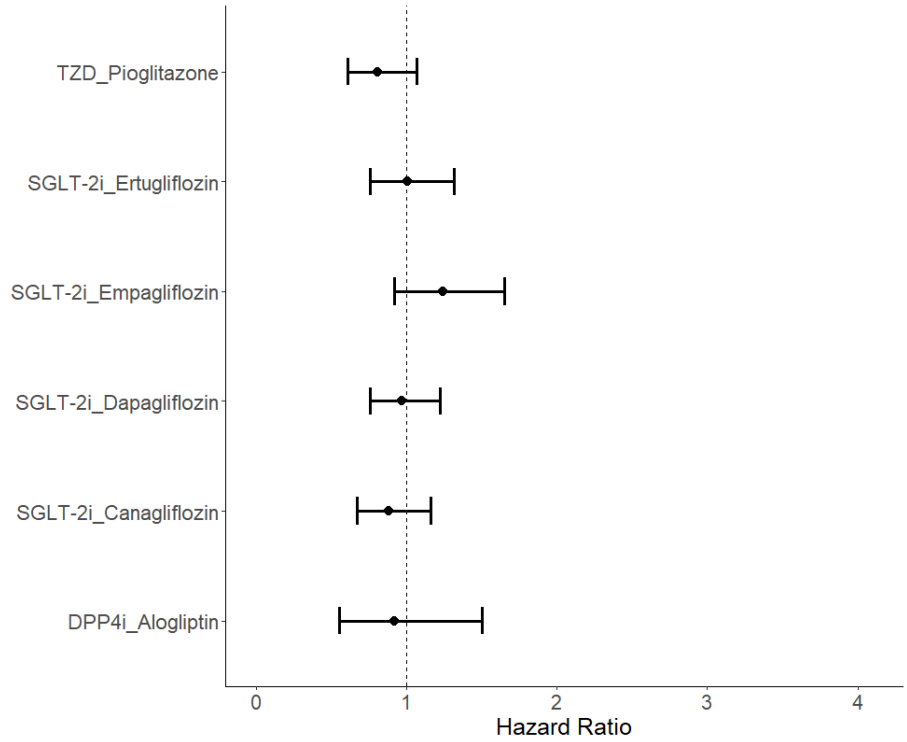

Figure 40. Hazard ratio for non-fatal stroke, relative to placebo, in the population with T2D and AS-CVD. The effect estimate for glimepiride could not be precisely estimated, so is not plotted.

Subpopulation with AS-CVD: non-fatal myocardial infarction (MI)

This analysis included eleven trials of ten treatments (Figure 41). Six studies reported hazard ratios, with a further five reporting the number of events. These events were rare in most trials, with rates equivalent to 0.4% to 9% of study participants experiencing non-fatal MI.

There was clear evidence for a decreased hazard of non-fatal MI compared to placebo for canagliflozin (Table 23). There was weaker evidence for a decreased hazard of non-fatal MI when compared with placebo for two other SGLT2-inhibitors, dapagliflozin and empagliflozin, as well as pioglitazone and sitagliptin. The treatment effect for vildagliptin was not estimable: the evidence for vildagliptin was drawn from a single small study, Phrommintikul 2019, which reported one event (1/25) on the dapagliflozin arm and zero events (0.24) on the vildagliptin arm. The estimated treatment effect for glimepiride was very uncertain: the evidence for glimepiride was drawn from a single study, Nissen 2008, which reported small event counts for both glimepiride (4/273) and pioglitazone (2/270) arms. There was strong evidence of a decreased hazard of non-fatal MI for canagliflozin compared to alogliptin, with a hazard ratio of 0.73 (95% CrI: 0.54, 0.99) (for the full table of active-active comparisons, see tab ‘Treatment Direct Effects’ in RQ1.2 results file CVD MI).

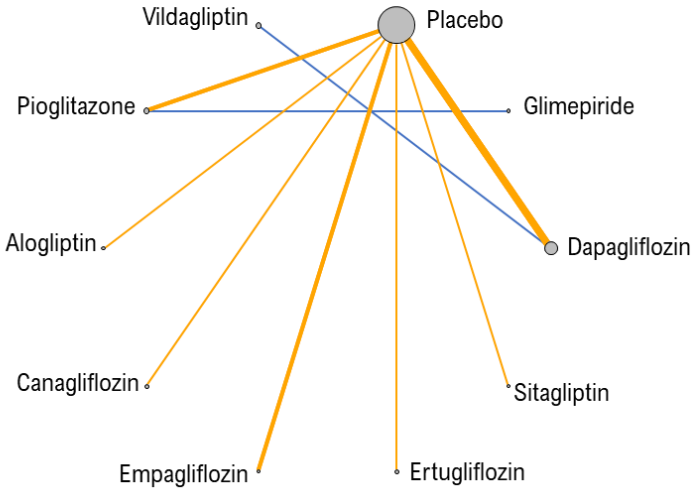

Figure 41. Network of evidence for non-fatal MI, in the subpopulation of people with diabetes and AS-CVD. Network edges where data were sparse (zero events on one study arm) are shown in blue.

Table 23. Hazard ratios for non-fatal MI in the subpopulation with diabetes and AS-CVD, relative to placebo. Median and range of rankings are also presented - these are less easy to interpret where hazard ratios are uncertain.

Class	Treatment	Hazard ratio Median (95% CrI)	Rank Median (range)
DPP4i	Alogliptin	1.08 (0.88, 1.33)	9 (4, 10)
	Sitagliptin	0.96 (0.81, 1.13)	6 (3, 9)
	Vildagliptin	*Not estimable*	*Not estimable*
Sulfonylurea	Glimepiride	1.84 (0.33, 14.61)	10 (2, 10)
SGLT-2i	Canagliflozin	0.79 (0.63, 0.99)	3 (1, 7)
	Dapagliflozin	0.88 (0.75, 1.03)	4 (2, 8)
	Empagliflozin	0.87 (0.70, 1.08)	4 (2, 9)
	Ertugliflozin	1.04 (0.86, 1.26)	8 (4, 10)
TZD	Pioglitazone	0.84 (0.66, 1.07)	4 (2, 8)

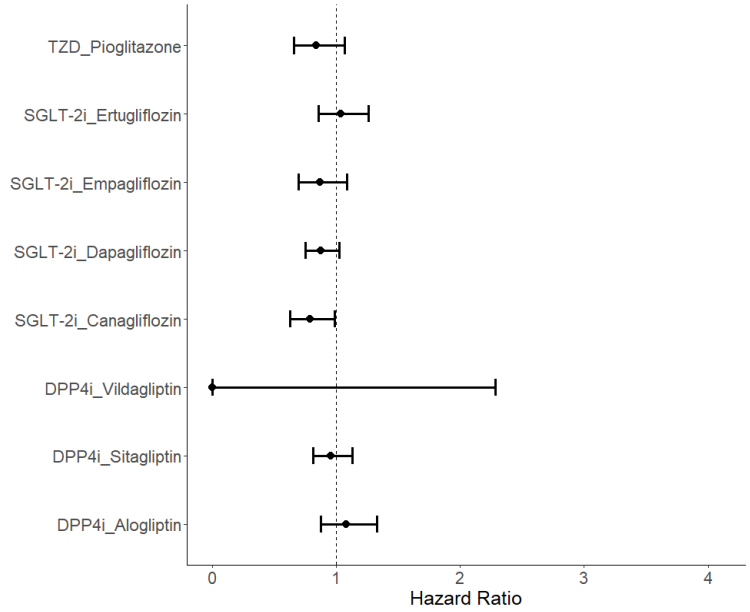

Figure 42. Hazard ratio for non-fatal MI, relative to placebo, in the population with T2D and AS-CVD. The treatment effect of glimepiride could not be precisely estimated, so is not plotted.

Subpopulation with AS-CVD: unstable angina

This analysis included seven trials of seven treatments (Figure 43). Three studies reported hazard ratios, with a further four reporting the number of events. Unstable angina was rare in all trials, developing in 0.3% to 6% of study participants.

The evidence for all treatments was uncertain, with the 95% credible intervals for these treatments including the probability of both protective and harmful effects relative to placebo (Table 24). The median treatment effects for dapagliflozin and ertugliflozin suggest weak evidence of lower hazard of unstable angina on these treatments. No clear differences were noted between active treatments (for the full table of active-active comparisons, see tab 'Treatment Direct Effects' in RQ1.2 results file CVD UA).

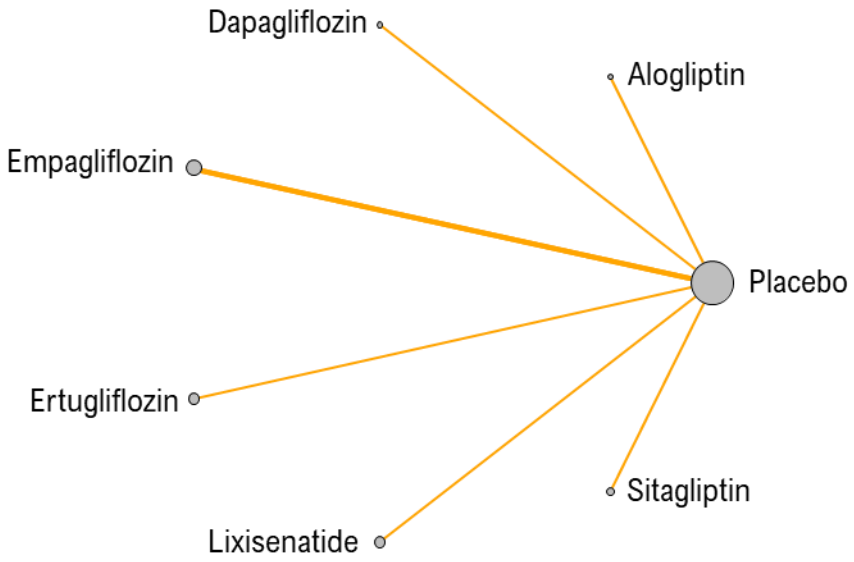

Figure 43. Network of evidence for unstable angina in the subpopulation of people with diabetes and AS-CVD.

Table 24. Hazard ratios for unstable angina in the subpopulation with diabetes and AS-CVD, relative to placebo. Posterior medians and ranges of rankings are also presented - these are less easy to interpret where hazard ratios are uncertain.

Class	Treatment	Hazard ratio Median (95% CrI)	Rank Median (range)
DPP4i	Alogliptin	0.90 (0.60, 1.37)	4 (1, 7)
	Sitagliptin	0.90 (0.70, 1.16)	4 (2, 7)
GLP-1RA	Lixisenatide	1.11 (0.47, 2.63)	6 (1, 7)
SGLT-2i	Dapagliflozin	0.40 (0.08, 1.50)	1 (1, 7)
	Empagliflozin	0.97 (0.73, 1.30)	5 (2, 7)
	Ertugliflozin	0.81 (0.63, 1.07)	3 (1, 6)

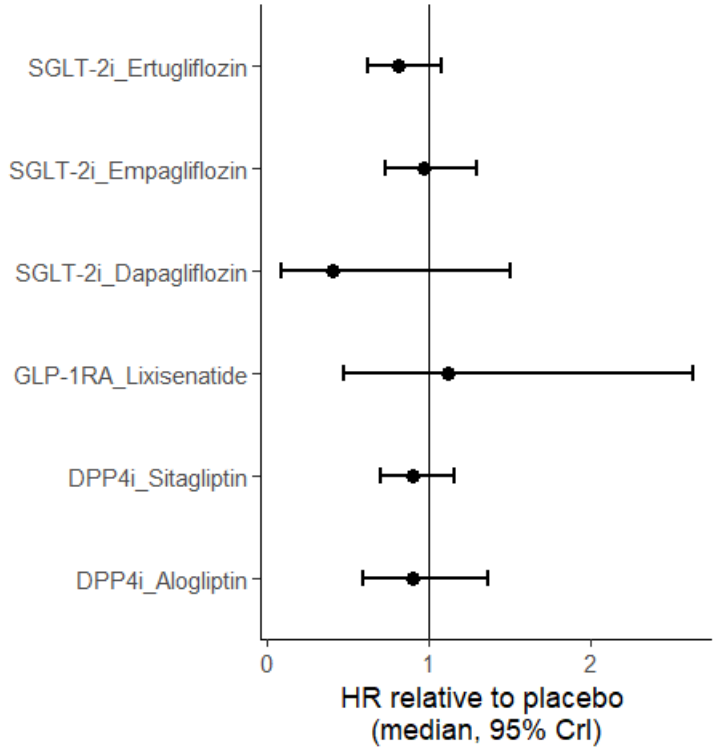

Figure 44. Hazard ratio for unstable angina, relative to placebo, in the population with T2D and AS-CVD.

Tables of model fit

Total residual deviance and DIC values were calculated within WinBUGS (version 1.4.3), using TSD standard code to estimate the total residual deviance, which is equivalent to the number of data points in a well-fitting model, and the inbuilt DIC tool to calculate pD and DIC, which is a measure of model fit penalised for complexity: i.e., the sum of posterior mean residual deviance (D-bar) and the number of effective parameters (pD).

Subpopulation with diabetes and CKD

Table 25. Model fit statistics for network meta-analysis (NMA) and unrelated mean-effects (UME) models of change in % HbA1c in the subpopulation with diabetes and CKD. [1]Lower values for model fit and DIC preferred.

Model	NMA effect structure	Number of data points	Model fit (total residual deviance[1])	pD (number of effective parameters)	DIC (penalised deviance[1])	Between-study SD median, (95% CrI)
NMA	Fixed	50	54.4	35.0	-28.5	-
	Random	50	50.2	39.8	-27.9	0.081 (0.005, 0.21)
UME	Fixed	50	48.4	38.0	-31.5	-

Table 26. Model fit statistics for network meta-analysis models of proportional change in weight (kg) in the subpopulation with diabetes and CKD. [1]Lower values for model fit and DIC preferred.

Model	NMA effect structure	Number of data points	Model fit (total residual deviance[1])	pD (number of effective parameters)	DIC (penalised deviance[1])	Between-study SD median, (95% CrI)
NMA (proportional change)	Fixed	18	18.8	12.9	39.4	-
	Random	18	*Model showed poor convergence*			
NMA (additive change)	Fixed	38	42.4	26.0	82.8	-
	Random	38	40.4	29.3	84.0	0.27 (0.01, 0.98)

Table 27. Model fit statistics for network meta-analysis models of CV mortality in the subpopulation with diabetes and CKD. [1]Lower values for model fit and DIC preferred.

Model	NMA effect structure	Number of data points	Model fit (total residual deviance[1])	pD (number of effective parameters)	DIC (penalised deviance[1])	Between-study SD median, (95% CrI)
NMA	Fixed	17	16.5	7.7	56.8	-
	Random	17	16.9	8.8	58.3	0.56 (0.03, 2.99)

Table 28. Model fit statistics for network meta-analysis models of hospitalisation for heart failure in the subpopulation with diabetes and CKD. [1]Lower values for model fit and DIC preferred.

Model	NMA effect structure	Number of data points	Model fit (total residual deviance[1])	pD (number of effective parameters)	DIC (penalised deviance[1])	Between-study SD median, (95% CrI)
NMA	Fixed	16	13.8	10.3	48.3	-
	Random	16	14.1	10.9	49.1	1.16 (0.05, 4.45)

Table 29. Model fit statistics for network meta-analysis models of three-point MACE in the subpopulation with diabetes and CKD. [1]Lower values for model fit and DIC preferred.

Model	NMA effect structure	Number of data points	Model fit (total residual deviance[1])	pD (number of effective parameters)	DIC (penalised deviance[1])	Between-study SD median, (95% CrI)
NMA	Fixed	4	6.8	3.0	23.0	-
	Random	4	4.1	4.0	21.4	1.78 (0.15, 4.78)

Table 30. Model fit statistics for network meta-analysis models of non-fatal stroke in the subpopulation with diabetes and CKD. [1]Lower values for model fit and DIC preferred.

Model	NMA effect structure	Number of data points	Model fit (total residual deviance[1])	pD (number of effective parameters)	DIC (penalised deviance[1])	Between-study SD median, (95% CrI)
NMA	Fixed	6	6.0	2.4	16.4	-

	Random	6				*Not estimable*

Table 31. Model fit statistics for network meta-analysis models of non-fatal MI in the subpopulation with diabetes and CKD. [1]Lower values for model fit and DIC preferred.

Model	NMA effect structure	Number of data points	Model fit (total residual deviance[1])	pD (number of effective parameters)	DIC (penalised deviance[1])	Between-study SD median, (95% CrI)
NMA	Fixed	7	7.3	5.3	27.9	-
	Random	7	7.4	5.3	27.9	1.89 (0.09, 4.78)

Table 32. Model fit statistics for network meta-analysis models of ESKD in the subpopulation with diabetes and CKD. [1]Lower values for model fit and DIC preferred.

Model	NMA effect structure	Number of data points	Model fit (total residual deviance[1])	pD (number of effective parameters)	DIC (penalised deviance[1])	Between-study SD median, (95% CrI)
NMA	Fixed	6	4.4	4.0	15.1	-
	Random	6				*Not estimable*

Subpopulation with diabetes and HF

Table 33. Model fit statistics for network meta-analysis (NMA) models of change in % HbA1c in the subpopulation with diabetes and HF. [1]Lower values for model fit and DIC preferred.

Model	NMA effect structure	Number of data points	Model fit (total residual deviance[1])	pD (number of effective parameters)	DIC (penalised deviance[1])	Between-study SD median, (95% CrI)
NMA	Fixed	5	5.0	5.0	9.0	-
	Random	5				*Not estimable*

Table 34. Model fit statistics for network meta-analysis models of CV mortality in the subpopulation with diabetes and HF. [1]Lower values for model fit and DIC preferred.

Model	NMA effect structure	Number of data points	Model fit (total residual deviance[1])	pD (number of effective parameters)	DIC (penalised deviance[1])	Between-study SD median, (95% CrI)
NMA	Fixed	7	7.0	6.9	11.9	-
	Random	7	*Not estimable*			

Table 35. Model fit statistics for network meta-analysis models of hospitalisation for heart failure in the subpopulation with diabetes and HF. [1]Lower values for model fit and DIC preferred.

Model	NMA effect structure	Number of data points	Model fit (residual deviance[1])	pD (number of effective parameters)	DIC (penalised deviance[1])	Between-study SD median, (95% CrI)
NMA	Fixed	11	11.5	10.0	14.3	-
	Random	11	11.1	10.9	14.8	0.86 (0.03, 4.59)

Table 36. Model fit statistics for network meta-analysis models of three-point MACE in the subpopulation with diabetes and HF. [1]Lower values for model fit and DIC preferred.

Model	NMA effect structure	Number of data points	Model fit (total residual deviance[1])	pD (number of effective parameters)	DIC (penalised deviance[1])	Between-study SD median, (95% CrI)
NMA	Fixed	8	7.3	7.0	-2.8	-
	Random	8	7.8	7.8	-1.4	0.76 (0.02, 4.56)

Table 37. Model fit statistics for network meta-analysis models of non-fatal stroke in the subpopulation with diabetes and HF. [1]Lower values for model fit and DIC preferred.

Model	NMA effect structure	Number of data points	Model fit (total residual deviance[1])	pD (number of effective parameters)	DIC (penalised deviance[1])	Between-study SD median, (95% CrI)
NMA	Fixed	6	9.1	0.7	21.0	-
	Random	6	*Not estimable*			

Table 38. Model fit statistics for network meta-analysis models of non-fatal MI in the subpopulation with diabetes and HF. [1]Lower values for model fit and DIC preferred.

Model	NMA effect structure	Number of data points	Model fit (total residual deviance[1])	pD (number of effective parameters)	DIC (penalised deviance[1])	Between-study SD median, (95% CrI)
NMA	Fixed	4	4.0	4.0	16.7	-
	Random	4	*Not estimable*			

Subpopulation with diabetes and AS-CVD

Table 39. Model fit statistics for network meta-analysis (NMA) and unrelated mean-effects (UME) models of change in % HbA1c in the subpopulation with diabetes and AS-CVD. [1]Lower values for model fit and DIC preferred.

Model	NMA effect structure	Number of data points	Model fit (total residual deviance[1])	pD (number of effective parameters)	DIC (penalised deviance[1])	Between-study SD median, (95% CrI)
NMA	Fixed	32	80.7	24.0	19.2	-
	Random	32	32.2	30.3	-23.0	0.31 (0.16, 0.69)
UME	Fixed	32	77.8	26.0	18.3	-
	Random	32	31.5	30.9	-23.1	0.30 (0.15, 0.74)

Table 40. Model fit statistics for network meta-analysis models of proportional change in weight (kg) in the subpopulation with diabetes and AS-CVD. [1]Lower values for model fit and DIC preferred.

Model	NMA effect structure	Number of data points	Model fit (total residual deviance[1])	pD (number of effective parameters)	DIC (penalised deviance[1])	Between-study SD median, (95% CrI)
NMA	Fixed	10	17.5	13.0	49.0	-
	Random	10	15.8	9.7	104.3	2.49 (0.14, 4.87)

Table 41. Model fit statistics for network meta-analysis models of CV mortality in the subpopulation with diabetes and AS-CVD. [1]Lower values for model fit and DIC preferred.

Model	NMA effect structure	Number of data points	Model fit (total residual deviance[1])	pD (number of effective parameters)	DIC (penalised deviance[1])	Between-study SD median, (95% CrI)
NMA	Fixed	16	13.9	12.8	25.4	-
	Random	16	14.6	13.7	27.0	0.50 (0.02, 2.73)

Table 42. Model fit statistics for network meta-analysis models of hospitalisation for heart failure in the subpopulation with diabetes and AS-CVD. [1]Lower values for model fit and DIC preferred.

Model	NMA effect structure	Number of data points	Model fit (total residual deviance[1])	pD (number of effective parameters)	DIC (penalised deviance[1])	Between-study SD median, (95% CrI)
NMA	Fixed	15	15.9	2.7	31.4	-
	Random	15	18.6	6.7	38.0	1.41 (0.03, 4.70)

Table 43. Model fit statistics for network meta-analysis models of three-point MACE in the subpopulation with diabetes and AS-CVD. [1]Lower values for model fit and DIC preferred.

Model	NMA effect structure	Number of data points	Model fit (total residual deviance[1])	pD (number of effective parameters)	DIC (penalised deviance[1])	Between-study SD median, (95% CrI)
NMA	Fixed	15	15.0	14.9	35.7	-

	Random	15				*Not estimable*

Table 44. Model fit statistics for network meta-analysis models of non-fatal stroke in the subpopulation with diabetes and AS-CVD. [1]Lower values for model fit and DIC preferred.

Model	NMA effect structure	Number of data points	Model fit (total residual deviance[1])	pD (number of effective parameters)	DIC (penalised deviance[1])	Between-study SD median, (95% CrI)
NMA	Fixed	10	8.7	7.6	10.9	-
	Random	10				*Not estimable*

Table 45. Model fit statistics for network meta-analysis models of non-fatal MI in the subpopulation with diabetes and AS-CVD. [1]Lower values for model fit and DIC preferred.

Model	NMA effect structure	Number of data points	Model fit (total residual deviance[1])	pD (number of effective parameters)	DIC (penalised deviance[1])	Between-study SD median, (95% CrI)
NMA	Fixed	16	15.5	12.9	44.8	-
	Random	16	15.2	13.9	45.5	0.96 (0.04, 4.06)

Table 46. Model fit statistics for network meta-analysis models of unstable angina in the subpopulation with diabetes and AS-CVD. [1]Lower values for model fit and DIC preferred.

Model	NMA effect structure	Number of data points	Model fit (total residual deviance[1])	pD (number of effective parameters)	DIC (penalised deviance[1])	Between-study SD median, (95% CrI)
NMA	Fixed	11	10.6	9.9	53.4	-
	Random	11	10.9	10.6	54.3	1.26 (0.05, 4.68)

References

- Balk EM, Earley A, Patel K, Trikalinos TA, Dahabreh IJ. Empirical assessment of within-arm correlation imputation in trials of continuous outcomes. Rockville: Agency for Healthcare Research and Quality; 2012.

- Daly C, Welton, S.J., Dias, S., Anwer,S., Ades, A.E. Meta-Analysis of Continuous Outcomes. Guideline Methodology Document 2: NICE Guidelines Technical Support Unit 2021.

- Dias S, Welton N, Sutton A, Ades A. NICE DSU Technical Support Document 2: A Generalised Linear Modelling Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials, 2011, last updated September 2016, available from https://www.sheffield.ac.uk/nice-dsu/tsds

- Dias S, Welton NJ, Sutton AJ, Caldwell DM, Lu G, Ades AE. Evidence Synthesis for Decision Making 4: Inconsistency in Networks of Evidence Based on Randomized Controlled Trials. Medical Decision Making 2013 33:641-656

WinBUGS code

Based on TSD2: Normal likelihood, identity link, additive FE treatment effects

```
#This code is part of
#Dias, S., Welton, N.J., Sutton, A.J. & Ades, A.E.
#NICE DSU Technical Support Document 2: A Generalised Linear Modelling
#Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials. 2011;
#last updated September 2016 (available #from http://www.nicedsu.org.uk).
#This work should be cited whenever the code is used whether in its standard form or adapted.

## This is adapted for the T2D medicines update

# Normal likelihood, identity link
# Arm and Trial-level data (treatment differences)
# FIXED effects model for multi-arm trials
# Added code to accept CFB, baseline-final and final values

model{                          # *** PROGRAM STARTS
for(i in 1:ns.a){               # LOOP THROUGH STUDIES WITH ARM DATA
   mu[i] ~ dnorm(0,.0001)         # vague priors for all trial baselines
   for (k in 1:na[i]) {           #  LOOP THROUGH ARMS
      delta[i,k] <- mu[i] + d[t[i,k]] - d[t[i,1]]  # FE
    }
#  summed residual deviance contribution for this trial
   resdev[i] <- sum(dev[i,1:na[i]])
 }

### SECTION A - estimation specific to studies reporting CFB
for(i in cb[1]:cb[2]){                          #   LOOP THROUGH STUDIES (CFB)
          for (k in 1:na[i]) {          #  LOOP THROUGH ARMS

# SE and precision for CFB
                    CFB_se[i,k] <- se[i,k]
                    prec[i,k] <- pow(CFB_se[i,k],-2)    ## precision

# Outcome measure: change from baseline (requires baseline)
                    yc[i,k] <- y[i,k]
                    phi[i,k] <- delta[i,k]

# Likelihood: univariate Normal
                    yc[i,k] ~ dnorm(phi[i,k], prec[i,k])

# Deviance: contribution for CFB means
      dev[i,k] <- (yc[i,k]-phi[i, k])*(yc[i,k]-phi[i, k])*prec[i,k]
     }              # END ARM LOOP
}  # END STUDY LOOP FOR CFB DATA
```

```
1
2   ### SECTION C - estimation specific to studies reporting follow-up values
3   for(i in pt[1]:pt[2]){        ##   LOOP THROUGH STUDIES (follow-up data)
4             for (k in 1:na[i]) {           ##  LOOP THROUGH ARMS
5   # SE
6                         f_se[i,k] <- se[i,k]
7                         prec[i,k] <- pow(f_se[i,k],-2)   ## precision
8
9   # Outcome measure: post-treatment mean
10                        ypt[i,k] <- y[i,k]
11                        phi[i,k] <- delta[i,k]
12
13  # Likelihood: univariate Normal
14                        ypt[i,k] ~ dnorm(phi[i,k], prec[i,k])
15
16  # Deviance: contribution for post-treatment mean
17        dev[i,k] <- (ypt[i,k]-phi[i, k])*(ypt[i,k]-phi[i, k])*prec[i,k]
18       }                # END ARM LOOP
19
20  }           ## END STUDY LOOP (follow-up data)
21
22
23  for(i in 1:ns2) {                 # LOOP THROUGH 2-ARM CONTRAST STUDIES
24     md[i,2] ~ dnorm(delta[(i+ns.a),2],prec.c[i,2]) # normal likel. 2-arm trials
25  #Deviance contribution for trial i (2-armed trials)
26     resdev[(i+ns.a)] <- (md[i,2]-delta[(i+ns.a),2])*(md[i,2]-delta[(i+ns.a),2])*prec.c[i,2]
27    }
28  for(i in (ns2+1):(ns2+ns3)) {       # LOOP THROUGH THREE-ARM CONTRAST STUDIES
29     for (k in 1:2) {    # set variance-covariance matrix
30        for (j in 1:2) {
31           Sigma[i,j,k] <- V[i]*(1-equals(j,k)) + prec.c[i,k+1]*equals(j,k)
32          }
33       }
34     Omega[i,1:2,1:2] <- inverse(Sigma[i,,])  #Precision matrix
35  # normal likelihood for 3-arm trials
36     md[i,2:3] ~ dmnorm(delta[(i+ns.a),2:3],Omega[i,1:2,1:2] )
37
38  #Deviance contribution for trial i
39     for(k in 1:2) {  # multiply vector & matrix
40        ydiff[i,k]<- md[i,(k+1)] - delta[(i+ns.a),(k+1)]
41        z[i,k]<- inprod2(Omega[i,k,1:2], ydiff[i,1:2])
42                        }
43     resdev[(i+ns.a)]<- inprod2(ydiff[i,1:2], z[i,1:2])
44    }
45
46  for(i in (ns2+ns3+1):ns.t) {       # LOOP THROUGH FOUR-ARM CONTRAST STUDIES
47     for (k in 1:3) {    # set variance-covariance matrix
```

```
        for (j in 1:3) {
          Sigma2[i,j,k] <- V[i]*(1-equals(j,k)) + prec.c[i,k+1]*equals(j,k)
        }
      }
    Omega2[i,1:3,1:3] <- inverse(Sigma2[i,,])  #Precision matrix
# normal likelihood for 4-arm trials
    md[i,2:na.c[i]] ~ dmnorm(delta[(i+ns.a),2:4],Omega2[i,1:3,1:3] )

#Deviance contribution for trial i
    for(k in 1:3) {  # multiply vector & matrix
      ydiff[i,k]<- md[i,(k+1)] - delta[(i+ns.a),(k+1)]
      z[i,k]<- inprod2(Omega2[i,k,1:3], ydiff[i,1:3])
                    }
    resdev[(i+ns.a)]<- inprod2(ydiff[i,1:3], z[i,1:3])
  }
for(i in 1:ns.t){              #   LOOP THROUGH ALL CONTRAST STUDIES
    delta[(i+ns.a),1] <- 0   # treatment effect is zero for control arm
    for (k in 2:na.c[i]) {          #  LOOP THROUGH ARMS
      var.c[i,k] <- pow(se.c[i,k],2)   # calculate variances
      prec.c[i,k] <- 1/var.c[i,k]     # set precisions
      delta[(i+ns.a),k] <-  d[t.c[i,k]] - d[t.c[i,1]]
     }
#bse[i,1] <- base_sd[i,1] / sqrt(n.c[i,1])
V[i] <- pow(base_sd[i,1],2)
  }

# Provide estimates of treatment effects T[k] on the natural scale
# Given a Mean Effect, meanA, for 'standard' treatment A,
# with precision (1/variance) precA
A ~ dnorm(meanA, precA)
seA <- sdA/sqrt(n_A)     ## SE
precA <- pow(seA,-2)   ## precision

for (k in 1:nt) { T[k] <- A + d[k]  }

## NO class effect
d[1]<-0      # treatment effect is zero for reference treatment / class

# vague priors for treatment effects within class
for (k in 2:nt){  d[k] ~ dnorm(0,.0001) }

tau <- pow(sd,-2)   # between-trial precision = (1/between-trial variance)
sd ~ dunif(0,sdUpper)     # vague prior for between-trial SD
totresdev <- sum(resdev[])          #Total Residual Deviance

# rank interventions
for (k in 1:nt) {
```

```
  # rk[k] <- nt+1-rank(d[],k)          # assumes positive diffs are  good
  rk[k] <- rank(d[],k)              # assumes negative diffs are  good
  best[k] <- equals(rk[k],1)       #calculate probability that treat k is best
  for (h in 1:nt) {
                    prob[h,k] <- equals(rk[k],h)                    # prob k is h-th best
          }
}

# MDs for all possible pair-wise comparisons - trts
for (c in 1:(nt-1)) {
          for (k in (c+1):nt) {
     diff[c,k] <- d[k] - d[c]
  }
}

#Stop unused variables causing error message

dv[1] <- n[1,1]
dv[2] <- base_m[1,1] + base_sd[51,1] + base_n[1,1]
dv[3] <- n.c[1,1]
dv[4] <- base_m.c[1,1] + base_sd.c[1,2] + base_n.c[1,2]

}                          # *** PROGRAM ENDS


Based on TSD2: Normal likelihood, identity link, additive RE treatment effects
#This code is part of
#Dias, S., Welton, N.J., Sutton, A.J. & Ades, A.E.
#NICE DSU Technical Support Document 2: A Generalised Linear Modelling
#Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials. 2011;
#last updated September 2016 (available #from http://www.nicedsu.org.uk).
#This work should be cited whenever the code is used whether in its standard form or adapted.

## This is adapted for the T2D medicines update

# Normal likelihood, identity link
# Arm and Trial-level data (treatment differences)
# RANDOM effects model for multi-arm trials
# Added code to accept CFB, baseline-final and final values (DONE)
# Altered the Norm_diff section to explicitly model 4-armed trials

model{                    # *** PROGRAM STARTS
for(i in 1:ns.a){           # LOOP THROUGH STUDIES WITH ARM DATA
   w.a[i,1] <- 0    # adjustment for multi-arm trials is zero for control arm
   delta[i,1] <- 0          # treatment effect is zero for control arm
   mu[i] ~ dnorm(0,.0001)         # vague priors for all trial baselines
```

```
    for (k in 1:na[i]) {          #  LOOP THROUGH ARMS
      theta[i,k] <- mu[i] + delta[i,k]  # model for linear predictor
     }
#  summed residual deviance contribution for this trial
    resdev[i] <- sum(dev[i,1:na[i]])
    for (k in 2:na[i]) {          # LOOP THROUGH ARMS
# trial-specific LOR distributions
      delta[i,k] ~ dnorm(MD[i,k],taud[i,k])
# mean of LOR distributions, with multi-arm trial correction
      MD[i,k] <-  d[t[i,k]] - d[t[i,1]] + sw.a[i,k]
# precision of LOR distributions (with multi-arm trial correction)
      taud[i,k] <- tau *2*(k-1)/k
# adjustment, multi-arm RCTs
      w.a[i,k] <- (delta[i,k] - d[t[i,k]] + d[t[i,1]])
# cumulative adjustment for multi-arm trials
      sw.a[i,k] <- sum(w.a[i,1:k-1])/(k-1)
    }
  }
### SECTION A - estimation specific to studies reporting CFB
for(i in cb[1]:cb[2]){                    #   LOOP THROUGH STUDIES (CFB)
          for (k in 1:na[i]) {          #  LOOP THROUGH ARMS

# SE and precision for CFB
                    CFB_se[i,k] <- se[i,k]
                    prec[i,k] <- pow(CFB_se[i,k],-2)    ## precision

# Outcome measure: change from baseline (requires baseline)
                    yc[i,k] <- y[i,k]
                    phi[i,k] <- theta[i,k]

# Likelihood: univariate Normal
                    yc[i,k] ~ dnorm(phi[i,k], prec[i,k])

# Deviance: contribution for CFB means
      dev[i,k] <- (yc[i,k]-phi[i, k])*(yc[i,k]-phi[i, k])*prec[i,k]
    }               # END ARM LOOP
}  # END STUDY LOOP FOR CFB DATA


### SECTION C - estimation specific to studies reporting follow-up values
for(i in pt[1]:pt[2]){       ##   LOOP THROUGH STUDIES (follow-up data)
          for (k in 1:na[i]) {          ##  LOOP THROUGH ARMS
# SE
                    f_se[i,k] <- se[i,k]
                    prec[i,k] <- pow(f_se[i,k],-2)   ## precision

# Outcome measure: post-treatment mean
```

```
                    ypt[i,k] <- y[i,k]
                    phi[i,k] <- theta[i,k]

# Likelihood: univariate Normal
                    ypt[i,k] ~ dnorm(phi[i,k], prec[i,k])

# Deviance: contribution for post-treatment mean
      dev[i,k] <- (ypt[i,k]-phi[i, k])*(ypt[i,k]-phi[i, k])*prec[i,k]
    }               # END ARM LOOP

}         ## END STUDY LOOP (follow-up data)


for(i in 1:ns2) {               # LOOP THROUGH 2-ARM CONTRAST STUDIES
  md[i,2] ~ dnorm(delta[(i+ns.a),2],prec.c[i,2]) # normal likel. 2-arm trials
#Deviance contribution for trial i (2-armed trials)
  resdev[(i+ns.a)] <- (md[i,2]-delta[(i+ns.a),2])*(md[i,2]-delta[(i+ns.a),2])*prec.c[i,2]
 }
for(i in (ns2+1):(ns2+ns3)) {       # LOOP THROUGH THREE-ARM CONTRAST STUDIES
  for (k in 1:2) {    # set variance-covariance matrix
    for (j in 1:2) {
      Sigma[i,j,k] <- V[i]*(1-equals(j,k)) + prec.c[i,k+1]*equals(j,k)
     }
   }
  Omega[i,1:2,1:2] <- inverse(Sigma[i,,])  #Precision matrix
# normal likelihood for 3-arm trials
  md[i,2:3] ~ dmnorm(delta[(i+ns.a),2:3],Omega[i,1:2,1:2] )

#Deviance contribution for trial i
  for(k in 1:2) {  # multiply vector & matrix
    ydiff[i,k]<- md[i,(k+1)] - delta[(i+ns.a),(k+1)]
    z[i,k]<- inprod2(Omega[i,k,1:2], ydiff[i,1:2])
                  }
  resdev[(i+ns.a)]<- inprod2(ydiff[i,1:2], z[i,1:2])
 }

for(i in (ns2+ns3+1):ns.t) {       # LOOP THROUGH FOUR-ARM CONTRAST STUDIES
  for (k in 1:3) {    # set variance-covariance matrix
    for (j in 1:3) {
      Sigma2[i,j,k] <- V[i]*(1-equals(j,k)) + prec.c[i,k+1]*equals(j,k)
     }
   }
  Omega2[i,1:3,1:3] <- inverse(Sigma2[i,,])  #Precision matrix
# normal likelihood for 4-arm trials
  md[i,2:na.c[i]] ~ dmnorm(delta[(i+ns.a),2:4],Omega2[i,1:3,1:3] )

#Deviance contribution for trial i
```

```
    for(k in 1:3) {  # multiply vector & matrix
      ydiff[i,k]<- md[i,(k+1)] - delta[(i+ns.a),(k+1)]
      z[i,k]<- inprod2(Omega2[i,k,1:3], ydiff[i,1:3])
                    }
    resdev[(i+ns.a)]<- inprod2(ydiff[i,1:3], z[i,1:3])
  }
for(i in 1:ns.t){              #   LOOP THROUGH ALL CONTRAST STUDIES
    w.c[i,1] <- 0    # adjustment for multi-arm trials is zero for control arm
    delta[(i+ns.a),1] <- 0   # treatment effect is zero for control arm
    for (k in 2:na.c[i]) {          #  LOOP THROUGH ARMS
      var.c[i,k] <- pow(se.c[i,k],2)   # calculate variances
      prec.c[i,k] <- 1/var.c[i,k]     # set precisions
# trial-specific LOR distributions
      delta[i+ns.a,k] ~ dnorm(MD[i+ns.a,k],taud.c[i,k])
# mean of LOR distributions, with multi-arm trial correction
      MD[i+ns.a,k] <-  d[t.c[i,k]] - d[t.c[i,1]] + sw.c[i,k]
# precision of LOR distributions (with multi-arm trial correction)
      taud.c[i,k] <- tau *2*(k-1)/k
# adjustment, multi-arm RCTs
      w.c[i,k] <- (delta[i+ns.a,k] - d[t.c[i,k]] + d[t.c[i,1]])
# cumulative adjustment for multi-arm trials
      sw.c[i,k] <- sum(w.c[i,1:k-1])/(k-1)


    }
#bse[i,1] <- base_sd[i,1] / sqrt(n.c[i,1])
V[i] <- pow(base_sd.c[i,1],2)
  }

# Provide estimates of treatment effects T[k] on the natural scale
# Given a Mean Effect, meanA, for 'standard' treatment A,
# with precision (1/variance) precA
A ~ dnorm(meanA, precA)
seA <- sdA/sqrt(n_A)     ## SE
precA <- pow(seA,-2)   ## precision

for (k in 1:nt) { T[k] <- A + d[k]  }

## NO class effect
d[1]<-0      # treatment effect is zero for reference treatment / class

# vague priors for treatment effects within class
for (k in 2:nt){  d[k] ~ dnorm(0,.0001) }

tau <- pow(sd,-2)   # between-trial precision = (1/between-trial variance)
sd ~ dunif(0,sdUpper)     # vague prior for between-trial SD
totresdev <- sum(resdev[])          #Total Residual Deviance
```

```
1
2   # rank interventions
3   for (k in 1:nt) {
4     # rk[k] <- nt+1-rank(d[],k)          # assumes positive diffs are  good
5     rk[k] <- rank(d[],k)              # assumes negative diffs are  good
6     best[k] <- equals(rk[k],1)       #calculate probability that treat k is best
7     for (h in 1:nt) {
8                         prob[h,k] <- equals(rk[k],h)                              # prob k is h-th best
9               }
10  }
11
12  # MDs for all possible pair-wise comparisons - trts
13  for (c in 1:(nt-1)) {
14              for (k in (c+1):nt) {
15        diff[c,k] <- d[k] - d[c]
16    }
17  }
18
19  #Stop unused variables causing error message
20
21  dv[1] <- n[1,1] + baseA[1]  ## baseA appears in mr coding
22  dv[2] <- base_m[1,1] + base_sd[51,1] + base_n[1,1]
23  dv[3] <- n.c[1,1]
24  dv[4] <- base_m.c[1,1] + base_sd.c[1,2] + base_n.c[1,2]
25
26  }                               # *** PROGRAM ENDS
27
```

Based on TSD2: Normal likelihood, log link, proportional FE treatment effects

```
29  # This code is part of
30  # Dias, S., Welton, N.J., Sutton, A.J. & Ades, A.E.
31  # NICE DSU Technical Support Document 2: A Generalised Linear Modelling
32  # Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials. 2011;
33  # last updated September 2016 (available #from http://www.nicedsu.org.uk).
34  # This work should be cited whenever the code is used whether in its standard form or adapted.
35
36  ## This is adapted for the T2D medicines update
37
38  # Normal likelihood, log link (i.e. ROM)
39  # FIXED effects model for multi-arm trials
40  # Added code to accept CFB, baseline-final and final values (DONE)
41
42  model{                          # *** PROGRAM STARTS
43  for(i in 1:ns){                 # LOOP THROUGH STUDIES WITH ARM DATA
44     mu[i] ~ dnorm(0,0.001)          # vague priors for all trial baselines
45     for (k in 1:na[i]) {            #  LOOP THROUGH ARMS
46  # model for linear predictor
```

```
        theta[i,k] <- mu[i] + d[t[i,k]] - d[t[i,1]]
      }
#  summed residual deviance contribution for this trial
   resdev[i] <- sum(dev[i,1:na[i]])
 }

### SECTION A - estimation specific to studies reporting CFB
for(i in cb[1]:cb[2]){                          #   LOOP THROUGH STUDIES (CFB)
          for (k in 1:na[i]) {          #  LOOP THROUGH ARMS

# SE and precision for CFB
                    CFB_se[i,k] <-  SD[i,k]/sqrt(n[i,k])
                    prec[i,k] <- pow(CFB_se[i,k],-2)

# Outcome measure: change from baseline (requires baseline)
                    yc[i,k] <- y[i,k]
                    phiB[i,k] <- base_m[i,k]
                    phi[i,k] <- phiF[i,k] - phiB[i,k]                                   ## CFB
                    log(phiF[i,k]) <- log(phiB[i,k]) + theta[i,k]           ## Follow-up

# Likelihood: univariate Normal
                    yc[i,k] ~ dnorm(phi[i,k], prec[i,k])

# Deviance: contribution for CFB means
      dev[i,k] <- (yc[i,k]-phi[i, k])*(yc[i,k]-phi[i, k])*prec[i,k]
     }              # END ARM LOOP
}  # END STUDY LOOP FOR CFB DATA

## SECTION B - estimation specific to studies reporting baseline and follow-up values
for(i in pp[1]:pp[2] ) {              ##   LOOP THROUGH STUDIES (baseline and follow-up)
          for (k in 1:na[i]) {          ##  LOOP THROUGH ARMS
# SE and variances at baseline and follow-up
                    base_se[i,k] <- base_sd[i,k] / sqrt(base_n[i,k])
                    base_var[i,k] <- pow(base_se[i,k],2)
                    f_se[i,k] <- SD[i,k]/sqrt(n[i,k])
                    f_var[i,k] <- pow(f_se[i,k],2)
                    prec[i,k] <- 1/f_var[i,k]

# Outcome measure: baseline and post-treatment means standardised by study SD
                    yp[i,k,1] <- base_m[i,k]
                    yp[i,k,2] <- y[i,k]
                    phiPP[i,k,1] <- ( basephi[i,k] )
                    log(phiPP[i,k,2]) <- ( log(basephi[i,k]) + theta[i,k] )
                    basephi[i,k] ~ dnorm(80, 0.0001)

## Likelihood: bivariate Normal
                    yp[i,k,1:2] ~ dmnorm(phiPP[i,k,1:2], sigmaInv[i, k, 1:2, 1:2])
```

```
1
2   # Precision matrix for mvnorm
3               sigma[i, k, 1, 1] <- base_var[i,k]
4               sigma[i, k, 1, 2] <- ( corr * base_se[i,k]  * f_se[i,k] )
5               sigma[i, k, 2, 1]  <- ( corr * base_se[i,k]  *f_se[i,k] )
6               sigma[i, k, 2, 2]  <- f_var[i,k]
7               sigmaInv[i, k, 1:2, 1:2] <- inverse(sigma[i, k, 1:2, 1:2])
8
9   # Deviance: Mahalanobis distance for trial i (baseline and follow-up data)
10                      for (j in 1:2) {                ## n of dimensions of mvnorm (i.e. bivariate)
11                      res[i, k, j] <- yp[i, k, j] - phiPP[i, k, j]
12                      temp[i, k, j] <- inprod(sigmaInv[i, k, j, 1:2], res[i, k, 1:2])
13                      }
14              Msq[i,k] <- inprod(res[i, k, 1:2], temp[i, k, 1:2])
15              M[i,k] <- sqrt(Msq[i,k])
16              dev[i,k] <- Msq[i,k]
17              } ## END ARM LOOP
18
19  }   ## END STUDY LOOP (baseline and follow-up data)
20
21  ### SECTION C - estimation specific to studies reporting follow-up values ONLY
22  #for(i in pt[1]:pt[2]){      ##   LOOP THROUGH STUDIES (follow-up data)
23  #           for (k in 1:na[i]) {            ##  LOOP THROUGH ARMS
24  # SE
25  #                       f_se[i,k] <- f_sd[i,k]/sqrt(n[i,k])   ## SE
26  #                       prec[i,k] <- pow(f_se[i,k],-2)   ## precision
27
28  # Outcome measure: post-treatment mean
29  #                       ypt[i,k] <- f_mean[i,k]
30  #                       log(phi[i,k]) <- theta[i,k]
31
32  # Likelihood: univariate Normal
33  #                       ypt[i,k] ~ dnorm(phi[i,k], prec[i,k])
34
35  # Deviance: contribution for post-treatment mean
36    #     dev[i,k] <- (ypt[i,k]-phi[i, k])*(ypt[i,k]-phi[i, k])*prec[i,k]
37      # }               # END ARM LOOP
38
39  #}          ## END STUDY LOOP (follow-up data)
40
41  totresdev <- sum(resdev[])          #Total Residual Deviance
42
43  ## NO class effect
44  d[1]<-0       # treatment effect is zero for reference treatment / class
45
46  # vague priors for treatment effects within class
47  for (k in 2:nt){  d[k] ~ dnorm(0,.0001)
```

```
dExp[k] <- exp(d[k])
}


# Calculate all relative treatment differences, array [nt-1, nt]
for (c in 1:(nt-1)) {  #
   for (k in (c+1):nt)  {     diff[c,k] <- T[k,1] - T[c,1] }
 }

# Ranking on relative scale, length=nt
for (k in 1:nt) {
        rk[k] <- rank(d[],k)         ## assumes events are "bad" / negative d values are good
        best[k] <- equals(rk[k],1)           ## probability that treat k is best
        for (h in 1:nt){ prob[h,k] <- equals(rk[k],h) }         ## probability that treat k is h-th best
 }

for (k in 1:nt) {
## treatment effect as difference between CFB on treatment k and reference treatment (metformin)
        T[k,1] <- T[k,2] - (A[2] - A[1])

        ## Absolute effect treatment k is absolute effect on treatment 1 multiplied by RoM for
treatment k

        for (j in 1:2){
                        MD[k,j] <- T[k,j] - T[1,j]
                        RoM[k,j] <- T[k,j] / T[1,j]
                }
        }

### dummy variables so that same dataset may be used for all models
dum[1] <-  A[1]
dum[2] <- corr + pp[1]
## NOT USED HERE
tau <- pow(sd,-2)   # between-trial precision = (1/between-trial variance)
sd ~ dunif(0,5)     # vague prior for between-trial SD

}                       # *** PROGRAM ENDS


Based on TSD2: Normal likelihood, log link, proportional RE treatment effects
#This code is part of
#Dias, S., Welton, N.J., Sutton, A.J. & Ades, A.E.
# NICE DSU Technical Support Document 2: A Generalised Linear Modelling
#Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials. 2011;
#last updated September 2016 (available #from http://www.nicedsu.org.uk).
#This work should be cited whenever the code is used whether in its standard form or adapted.
```

```
1
2   ## This is adapted for the T2D medicines update
3
4   # Normal likelihood, identity link
5   # Random effects model for multi-arm trials
6   # Added code to accept CFB, baseline-final and final values (DONE)
7
8
9   model{                          # *** PROGRAM STARTS
10  for(i in 1:ns){                 # LOOP THROUGH STUDIES WITH ARM DATA
11     w.a[i,1] <- 0    # adjustment for multi-arm trials is zero for control arm
12     delta[i,1] <- 0           # treatment effect is zero for control arm
13     mu[i] ~ dnorm(0,0.001)          # vague priors for all trial baselines
14     for (k in 1:na[i]) {           #  LOOP THROUGH ARMS
15        theta[i,k] <- mu[i] + delta[i,k]  # model for linear predictor
16      }
17  #  summed residual deviance contribution for this trial
18     resdev[i] <- sum(dev[i,1:na[i]])
19     for (k in 2:na[i]) {           # LOOP THROUGH ARMS
20  # trial-specific LOR distributions
21        delta[i,k] ~ dnorm(md[i,k],taud.a[i,k])
22  # mean of LOR distributions, with multi-arm trial correction
23        md[i,k] <-  d[t[i,k]] - d[t[i,1]] + sw.a[i,k]
24  # precision of LOR distributions (with multi-arm trial correction)
25        taud.a[i,k] <- tau *2*(k-1)/k
26  # adjustment, multi-arm RCTs
27        w.a[i,k] <- (delta[i,k] - d[t[i,k]] + d[t[i,1]])
28  # cumulative adjustment for multi-arm trials
29        sw.a[i,k] <- sum(w.a[i,1:k-1])/(k-1)
30       }
31    }
32
33  ### SECTION A - estimation specific to studies reporting CFB
34  for(i in cb[1]:cb[2]){                          #   LOOP THROUGH STUDIES (CFB)
35            for (k in 1:na[i]) {           #  LOOP THROUGH ARMS
36
37  # SE and precision for CFB
38                      CFB_se[i,k] <- sqrt(n[i,k]) * SD[i,k]
39                      prec[i,k] <- pow(CFB_se[i,k],-2)
40
41  # Outcome measure: change from baseline (requires baseline)
42                      yc[i,k] <- y[i,k]
43                      phiB[i,k] <- base_m[i,k]
44                      phi[i,k] <- phiF[i,k] - phiB[i,k]                                   ## CFB
45                      log(phiF[i,k]) <- log(phiB[i,k]) + theta[i,k]              ## Follow-up
46
47  # Likelihood: univariate Normal
```

```
                    yc[i,k] ~ dnorm(phi[i,k], prec[i,k])

# Deviance: contribution for CFB means
     dev[i,k] <- (yc[i,k]-phi[i, k])*(yc[i,k]-phi[i, k])*prec[i,k]
    }               # END ARM LOOP
}  # END STUDY LOOP FOR CFB DATA

## SECTION B - estimation specific to studies reporting baseline and follow-up values
for(i in pp[1]:pp[2] ) {              ##   LOOP THROUGH STUDIES (baseline and follow-up)
          for (k in 1:na[i]) {        ##  LOOP THROUGH ARMS

# SE and variances at baseline and follow-up
                    base_se[i,k] <- base_sd[i,k]/sqrt(base_n[i,k])
                    base_var[i,k] <- pow(base_se[i,k],2)
                    f_se[i,k] <- SD[i,k]/sqrt(n[i,k])
                    f_var[i,k] <- pow(f_se[i,k],2)
                    prec[i,k] <- 1/f_var[i,k]

# Outcome measure: baseline and post-treatment means
                    yp[i,k,1] <- base_m[i,k]
                    yp[i,k,2] <- y[i,k]
                    phiPP[i,k,1] <- ( basephi[i,k] )
                    log(phiPP[i,k,2]) <- ( log(basephi[i,k]) + theta[i,k] )
                    basephi[i,k] ~ dnorm(90, 0.0001)

## Likelihood: bivariate Normal
                    yp[i,k,1:2] ~ dmnorm(phiPP[i,k,1:2], sigmaInv[i, k, 1:2, 1:2])

# Precision matrix for mvnorm
          sigma[i, k, 1, 1] <- base_var[i,k]
          sigma[i, k, 1, 2] <- ( corr * base_se[i,k]  * f_se[i,k] )
          sigma[i, k, 2, 1]  <- ( corr * base_se[i,k]  *f_se[i,k] )
          sigma[i, k, 2, 2]  <- f_var[i,k]
          sigmaInv[i, k, 1:2, 1:2] <- inverse(sigma[i, k, 1:2, 1:2])

# Deviance: Mahalanobis distance for trial i (baseline and follow-up data)
                    for (j in 1:2) {              ## n of dimensions of mvnorm (i.e. bivariate)
                    res[i, k, j] <- yp[i, k, j] - phiPP[i, k, j]
                    temp[i, k, j] <- inprod(sigmaInv[i, k, j, 1:2], res[i, k, 1:2])
                    }
           Msq[i,k] <- inprod(res[i, k, 1:2], temp[i, k, 1:2])
           M[i,k] <- sqrt(Msq[i,k])
          dev[i,k] <- Msq[i,k]
          } ## END ARM LOOP

}  ## END STUDY LOOP (baseline and follow-up data)
```

```
### SECTION C - estimation specific to studies reporting follow-up values
#for(i in pt[1]:pt[2]){      ##   LOOP THROUGH STUDIES (follow-up data)
#          for (k in 1:na[i]) {          ##  LOOP THROUGH ARMS
#
# SE
#                    f_se[i,k] <- f_sd[i,k]/sqrt(n[i,k])   ## SE
#                    prec[i,k] <- pow(f_se[i,k],-2)   ## precision

# Outcome measure: post-treatment mean
#                    ypt[i,k] <- f_mean[i,k]
#                    log(phi[i,k]) <- theta[i,k]

# Likelihood: univariate Normal
#                    ypt[i,k] ~ dnorm(phi[i,k], prec[i,k])

# Deviance: contribution for post-treatment mean
 #     dev[i,k] <- (ypt[i,k]-phi[i, k])*(ypt[i,k]-phi[i, k])*prec[i,k]
   # }              # END ARM LOOP

#}         ## END STUDY LOOP (follow-up data)


totresdev <- sum(resdev[])          #Total Residual Deviance

## NO class effect
d[1]<-0      # treatment effect is zero for reference treatment / class
dExp[1] <- 1

# vague priors for treatment effects within class
for (k in 2:nt){  d[k] ~ dnorm(0,.0001)
                                                                    dExp[k] <- exp(d[k])
}

tau <- pow(sd,-2)   # between-trial precision = (1/between-trial variance)
sd ~ dunif(0,5)     # vague prior for between-trial SD

# Calculate all relative treatment differences, array [nt-1, nt]
for (c in 1:(nt-1)) {  #
   for (k in (c+1):nt)  {     diff[c,k] <- T[k,1] - T[c,1] }
 }

# Ranking on relative scale, length=nt
for (k in 1:nt) {
          rk[k] <- rank(d[],k)          ## assumes events are "bad" / negative d values are good
          best[k] <- equals(rk[k],1)          ## probability that treat k is best
          for (h in 1:nt){ prob[h,k] <- equals(rk[k],h) }          ## probability that treat k is h-th best
 }
```

```
1
2   for (k in 1:nt) {
3   ## treatment effect as difference between CFB on treatment k and reference treatment (metformin)
4           T[k,1] <- T[k,2] - (A[2] - A[1])
5
6           ## Absolute effect treatment k is absolute effect on treatment 1 multiplied by RoM for
7   treatment k
8
9           for (j in 1:2){
10                          MD[k,j] <- T[k,j] - T[1,j]
11                          RoM[k,j] <- T[k,j] / T[1,j]
12                  }
13          }
14
15  }                       # *** PROGRAM ENDS
16  #
```

Based on TSD2: Shared parameter model of hazard rates with FE structure on treatment effects, data reported as either log HRs (with SE or 95% CI) or number of events.

```
20  model{                  # *** PROGRAM STARTS
21  for(i in 1:rs){                 ## LOOP THROUGH r/n STUDIES
22          for (k in 1:na[i]) {        # LOOP THROUGH ARMS
23                  r[i,k] ~ dbin(p[i,k],n[i,k])
24
25  # model for linear predictor
26          #       cloglog(p[i,k]) <- log(time[i]) + mu[i] + d[t[i,k]] - d[t[i,1]]
27  # cloglog truncated to avoid arithmetic overflow when close to 0 or 1
28  # see Ntzoufras(2009, Chapter 7)
29          cloglog(p[i,k]) <- eta[i,k]*(1-step(-xi1-eta[i,k]))*(1-step(eta[i,k]-xi2))
30  -xi1*step(-xi1-eta[i,k])+ xi2*step(eta[i,k]-xi2)
31                  eta[i,k] <- log(time[i]) + mu[i] + d[t[i,k]] - d[t[i,1]]
32                  rhat[i,k] <- p[i,k] * n[i,k]
33                  dev[i,k] <- 2 * (r[i,k] * (log(r[i,k])-log(rhat[i,k]))+ (n[i,k]-r[i,k]) * (log(n[i,k]-r[i,k]) -
34  log(n[i,k]-rhat[i,k])))
35          resdev[i] <- sum(dev[i,1:na[i]])
36  }                       ## END STUDY LOOP
37
38  for(i in rs+1:rs+cs2) {             # LOOP THROUGH 2-ARM CONTRAST STUDIES
39     y[i,2] ~ dnorm(theta[i,2],prec[i,2]) # normal likelihood for 2-arm trials
40          resdev[i] <- sum(dev[i,2:na[i]])
41          for(k in 2:na[i]) {
42                   y[i,k] <- log(HR[i,k])
43                  se[i,k] <- ( log(upperCI[i,k]) - log(lowerCI[i,k]) ) / 3.92
44          var[i,k] <- pow(se[i,k],2)   # calculate variances
45           prec[i,k] <- 1/var[i,k]     # set precisions
46                  dev[i,k] <- (y[i,k]-theta[i,k])*(y[i,k]-theta[i,k])*prec[i,k]
```

```
                    }          # Deviance contribution
 }


for(i in 1:rs+cs2+csp2){              #   LOOP THROUGH ALL STUDIES
   mu[i] ~ dnorm(0, 0.0001)


 for (k in 2:na[i]) {           #  LOOP THROUGH ARMS
## linear predictor
      theta[i,k] <-  d[t[i,k]] - d[t[i,1]]

    }
          time[i] <- t_months[i]
}

totresdev <- sum(resdev[])          #Total Residual Deviance
d[1]<-0      # treatment effect is zero for control arm
dExp[1]<-1      # treatment effect is zero for control arm

# cloglog truncation values
xi1 <- 10
xi2 <- 3

# vague priors for treatment effects
for (k in 2:nt){  d[k] ~ dnorm(0,.0001)
dExp[k] <- exp(d[k])        }
sd ~ dunif(0,5)     # vague prior for between-trial SD
tau <- pow(sd,-2)   # between-trial precision = (1/between-trial variance)

# Provide estimates of treatment effects T[k] on the natural scale
# Given a Mean Effect, meanA, for 'standard' treatment A,
# with precision (1/variance) precA
#A ~ dnorm(meanA,precA)
#for (k in 1:nt) { T[k] <- A + d[k]  }

for (c in 1:(nt-1)) {
          for (k in (c+1):nt) {
                    lhr[c,k] <- (d[k]-d[c])
                    log(hr[c,k]) <- lhr[c,k]              }
}

# rank interventions
for (k in 1:nt) {
  # rk[k] <- nt+1-rank(d[],k)          # assumes positive diffs are  good
  rk[k] <- rank(d[],k)             # assumes negative diffs are  good
  best[k] <- equals(rk[k],1)       #calculate probability that treat k is best
```

```
  for (h in 1:nt) {
                prob[h,k] <- equals(rk[k],h)                    # prob k is h-th best
        }
}

}                       # *** PROGRAM ENDS
```

Based on TSD2: Shared parameter model of hazard rates with RE structure on treatment effects, data reported as either log HRs (with SE or 95% CI) or number of events.

```
model{                  # *** PROGRAM STARTS
for(i in 1:rs){                 ## LOOP THROUGH r/n STUDIES
        for (k in 1:na[i]) {        # LOOP THROUGH ARMS
                r[i,k] ~ dbin(p[i,k],n[i,k])
# model for linear predictor
        #       cloglog(p[i,k]) <- log(time[i]) + mu[i] + delta[i,k]
# cloglog truncated to avoid arithmetic overflow when close to 0 or 1
# see Ntzoufras(2009, Chapter 7)
        cloglog(p[i,k]) <- eta[i,k]*(1-step(-xi1-eta[i,k]))*(1-step(eta[i,k]-xi2))
-xi1*step(-xi1-eta[i,k])+ xi2*step(eta[i,k]-xi2)
                eta[i,k] <- log(time[i]) + mu[i] + delta[i,k]
                rhat[i,k] <- p[i,k] * n[i,k]
                dev[i,k] <- 2 * (r[i,k] * (log(r[i,k])-log(rhat[i,k]))+ (n[i,k]-r[i,k]) * (log(n[i,k]-r[i,k]) -
log(n[i,k]-rhat[i,k])))
        resdev[i] <- sum(dev[i,1:na[i]])
}               ## END STUDY LOOP

for(i in rs+1:rs+cs2) {         # LOOP THROUGH 2-ARM CONTRAST STUDIES
   y[i,2] ~ dnorm(delta[i,2],prec[i,2]) # normal likelihood for 2-arm trials
        resdev[i] <- sum(dev[i,2:na[i]])

        for(k in 2:na[i]) {
                y[i,k] <- log(HR[i,k])
                se[i,k] <- ( log(upperCI[i,k]) - log(lowerCI[i,k]) ) / 3.92
        var[i,k] <- pow(se[i,k],2)   # calculate variances
         prec[i,k] <- 1/var[i,k]     # set precisions
                dev[i,k] <- (y[i,k]-delta[i,k])*(y[i,k]-delta[i,k])*prec[i,k]
                }       # Deviance contribution
 }


for(i in 1:rs+cs2+csp2){        #   LOOP THROUGH ALL STUDIES
   w[i,1] <- 0    # adjustment for multi-arm trials is zero for control arm
   delta[i,1] <- 0          # treatment effect is zero for control arm
   mu[i] ~ dnorm(0, 0.0001)
```

```
 for (k in 2:na[i]) {          #  LOOP THROUGH ARMS


# trial-specific LOR distributions
      delta[i,k] ~ dnorm(md[i,k],taud[i,k])
# mean of LOR distributions, with multi-arm trial correction
      md[i,k] <-  d[t[i,k]] - d[t[i,1]] + sw[i,k]
# precision of LOR distributions (with multi-arm trial correction)
      taud[i,k] <- tau *2*(k-1)/k
# adjustment, multi-arm RCTs
      w[i,k] <- (delta[i,k] - d[t[i,k]] + d[t[i,1]])
# cumulative adjustment for multi-arm trials
      sw[i,k] <- sum(w[i,1:k-1])/(k-1)
     }
          time[i] <- t_months[i]
}

totresdev <- sum(resdev[])          #Total Residual Deviance
d[1]<-0      # treatment effect is zero for control arm

# cloglog truncation values
xi1 <- 10
xi2 <- 3

# vague priors for treatment effects
for (k in 2:nt){  d[k] ~ dnorm(0,.0001) }
sd ~ dunif(0,5)     # vague prior for between-trial SD
tau <- pow(sd,-2)   # between-trial precision = (1/between-trial variance)

# Provide estimates of treatment effects T[k] on the natural scale
# Given a Mean Effect, meanA, for 'standard' treatment A,
# with precision (1/variance) precA
#A ~ dnorm(meanA,precA)
#for (k in 1:nt) { T[k] <- A + d[k]  }

for (c in 1:(nt-1)) {
          for (k in (c+1):nt) {
                    lhr[c,k] <- (d[k]-d[c])
                    log(hr[c,k]) <- lhr[c,k]              }
}

# rank interventions
for (k in 1:nt) {
  # rk[k] <- nt+1-rank(d[],k)          # assumes positive diffs are  good
  rk[k] <- rank(d[],k)             # assumes negative diffs are  good
  best[k] <- equals(rk[k],1)       #calculate probability that treat k is best
```

```
1       for (h in 1:nt) {
2                       prob[h,k] <- equals(rk[k],h)                    # prob k is h-th best
3               }
4       }
5
6       }                       # *** PROGRAM ENDS
7
8       END
```