National Institute for Health and Care Excellence

Final

Type 2 diabetes in adults: management

Evidence review A for SGLT2 inhibitors for people with chronic kidney disease and type 2 diabetes

NICE guideline 28

Evidence reviews underpinning recommendations 1.8.12 to 1.8.15
and research recommendations in the NICE guideline
24 November 2021

Final

These evidence reviews were developed by the NICE Guideline updates team

ISBN: 978-1-4731-1477-7

Contents

SGLT2 inhibitors for people with chronic kidney disease and type 2 diabetes

1.1 Review question

What is the clinical and cost effectiveness of SGLT2 inhibitors for children, young people and adults with chronic kidney disease and type 2 diabetes?

1.1.1 Introduction

NG203 chronic kidney disease in adults: assessment and management currently includes review evidence and recommendations on SGLT2 inhibitors for people with proteinuria and type 2 diabetes. This review will expand that to all people with chronic kidney disease (CKD) and type 2 diabetes and look at the cost-effectiveness of this intervention. The scope of the NICE guideline on CKD includes adults, children and young people and so the review question included these population groups. However, it was subsequently decided that the recommendations are better placed in the NICE guideline on type 2 diabetes, with a cross referral from the NICE guideline on CKD. The scope for this guideline is limited to adults, and recommendations arising from this review are for adults only. The evidence identified was all in adult population, and SGLT2 inhibitors are only currently licensed for people over the age of 18. At the time of publication (September 2021), the only SGLT2 inhibitors licensed for use in chronic kidney disease are canagliflozin and dapagliflozin.

1.1.2 Summary of the protocol

Table 1: PICO table for SGLT2 inhibitors for people with chronic kidney disease and type 2 diabetes

Population	Adults, children, and young people with CKD and type 2 diabetes **Exclusion:** • people receiving renal replacement therapy (RRT) • people with acute kidney injury combined with rapidly progressive glomerulonephritis • people receiving palliative care. • Pregnant women with type 2 diabetes. • People with type 1 diabetes. • People with type 2 diabetes who are hyperglycaemic and require rescue treatment. • Trials including a mixed population (for example of people with and without CKD or with and without type 2 diabetes) will be excluded unless a subgroup analysis for people with CKD and type 2 diabetes is reported, or this population comprises >85% of the trial population.
Intervention	Existing therapy + SGLT2 inhibitors (including canagliflozin, dapagliflozin, ertugliflozin, empagliflozin) • Examples of existing therapy could include antidiabetic medications, such as metformin, or therapy to treat CKD, such as an ARB or ACEi
Comparator	• Existing therapy • Existing therapy plus Placebo
Outcome	• Proteinuria (measured as urinary albumin to creatinine ratio, ACR)

	• CKD progression: occurrence of end stage kidney disease (End stage renal disease or end stage kidney disease as reported by the study - we will report doubling of serum creatinine, renal replacement therapy and transplant separately if data is available) • Mortality (all-cause and cardiovascular) • Specific morbidity (at the longest timepoint reported by the study): - advancement of renal bone disease, - vascular calcification, - cardiovascular impact, including macrovascular events (non-fatal MI, non-fatal stroke, hospitalisation for heart failure) - anaemia • health-related quality of life • estimated glomerular filtration rate (eGFR) (continuous outcome, or dichotomous outcome as number of participants with eGFR reduction>40% or 50%, as reported) • Adverse outcomes: - Acute kidney injury, - drug specific: hypotension/falls, hypoglycaemia, amputations, genitourinary infections, fractures, diabetic ketoacidosis

1.1.3 Methods and process

This evidence review was developed using the methods and process described in Developing NICE guidelines: the manual and the methods described in Appendix L. Methods specific to this review question are described below and in the review protocol in **Error! Reference source not found.**

Declarations of interest were recorded according to NICE's conflicts of interest policy

The following methods were specific for this review:

1. The evidence was analysed by SGLT2 inhibitors as a class of medication because it was assumed that medications within this class would have similar mechanisms of action and similar pharmaceutical effects. Specific medications are listed in the forest plots.
2. Mortality outcomes, cardiovascular outcomes, CKD progression and reduction in estimated glomerular filtration rate (eGFR) by 40% or 50% are reported as hazard ratios over the course of the study. Clinical decision thresholds used to rate imprecision for these outcomes were set at a HR of 1, meaning that any difference in effect was considered clinically important.
3. Other dichotomous outcomes are reported as risk ratios at the longest timepoint reported by the study. For these outcomes, clinical decision thresholds of 0.8 and 1.25 were used to rate imprecision.
4. The eGFR and urinary albumin to creatinine ratio (ACR) continuous outcomes are reported as mean differences at 6 months after treatment onset and for the longest available time point, providing it is over 2 years. A clinical decision threshold of 0.5 of the median standard deviations of the comparison group arms was used to rate imprecision (Norman et al. 2003).
5. The clinical decision thresholds used to rate imprecision were used as a starting point when judging the clinical importance of effects. However, the committee also took into account the total weight of evidence across outcomes and the uncertainty in the effect estimates. The clinical importance of effects is discussed in section 1.1.10 The committee's discussion and interpretation of the evidence.
6. Where trials report data for more than one dose, the dose within the BNF recommended range (as of May 2021, taking account correction for renal function) has been chosen for inclusion in the analysis.

7. Where possible, data has been stratified by:
    a. different albuminuria categories at baseline: A1 (ACR <3 mg/mmol) A2 (ACR 3-30 mg/mmol) and A3 (ACR >30 mg/mmol)
    b. different eGFR categories at baseline (eGFR 30-45,45-60 and >60 ml/min/1.73 m2)

    These categorisations were specified *a priori* in the review protocol and were chosen to align with the international Kidney Disease Improving Global Outcomes (KDIGO) classification scheme for eGFR and ACR. This scheme is recommended in the NICE guideline on chronic kidney disease and is used to guide the assessment and management of people with CKD.

8. Where there was significant (I^2>50%) heterogeneity a subgroup analysis by drug type was carried out.
9. Results were reported separately in the GRADE profiles only when there was evidence for a difference in effect across subgroups (test for subgroup differences, p<0.05).
10. When summarising the effectiveness evidence (see Table), hazard ratios and risk ratios below 1 (with 95% confidence intervals below 1) were described as an effect that favours treatment; mean differences (other than eGFR continuous) below 0 (with 95% confidence intervals below 0) were described as an effect that favours treatment. eGFR continuous outcomes below 0 were described as favouring placebo. Where 95% CIs crossed 1 (HR/RR) or 0 (MD) effects were described as being unable to differentiate between treatment and placebo.
11. Evidence was included from trials with broader populations where effects from a subgroup of participants matching the review protocol were reported. The committee agreed that eGFR<60 or ACR>3 mg/mol could be used as proxies for CKD, in line with the criteria for diagnosing CKD specified in the NICE guideline on CKD.

1.1.4 Effectiveness evidence

1.1.4.1 Included studies

A systematic search was carried out to identify randomised controlled trials (RCTs) and systematic reviews of RCTs (see appendix C for the literature search strategy). In total, 321 references were identified for screening at title and abstract level with 241 excluded at this level. Full texts were ordered to be screened for 80 references.

In total, 37 papers reporting 10 randomised controlled trials were included based on their relevance to the review protocol (**Error! Reference source not found.**). The clinical evidence study selection is presented as a PRISMA diagram in Appendix C. The populations of all included studies were adults with type 2 diabetes and CKD. No studies were included that reported the effectiveness of SGLT2 inhibitors in children and young people.

1.1.4.2 Excluded studies

Studies that were excluded at the full-text screening stage and reasons for exclusion can be found in Appendix J.

1.1.5 Summary of studies included in the effectiveness evidence.

Table 21: Summary of studies included in the effectiveness evidence.

Note that some of the included evidence was from subgroups of larger trials. The population and outcomes listed in the table below relate specifically to the subgroups that are reported in this evidence review, rather than the overall trial populations.

Trial [1], author, year, sample size	Population[2]	Follow-up	Baseline eGFR and ACR (mean/median/n%)	Intervention	Comparator	Outcome[2]
Adults with CKD and type 2 diabetes						
SGLT2 vs placebo						
Subgroup of VERTIS CV (Cherney 2021) n=1807	Adults with Type 2 Diabetes and CKD with eGFR 30-60 or ACR A2 and A3	3 years (median), 3.5 (mean)	CKD stage 3 subgroup: Mean eGFR 48.9 ml/min/1.73 m2 Median ACR 3.5mg/mmol	Ertugliflozin 5 or 15mg + existing therapy at study entry	Placebo + +existing therapy at study entry	Renal composite - doubling of baseline serum creatinine, kidney dialysis/transplant or renal death eGFR >2 years Percentage change from baseline ACR at last available data point
Subgroup of CANVAS (Neuen 2019) N=2039	Adults with Type 2 Diabetes and CKD with eGFR 30-60 or ACR A2 and A3	3.6 years (mean)	Subgroup eGFR 30-60: Mean eGFR 49.1 ml/min/1.73 m2 Median ACR 2.4 mg/mmol	Canagliflozin 100mg	Placebo	Renal composite (as above) Cardiovascular composite – (as above Cardiovascular death Fatal/non-fatal MI

Trial [1], author, year, sample size	Population[2]	Follow-up	Baseline eGFR and ACR (mean/median/n%)	Intervention	Comparator	Outcome[2]
						Fatal/non-fatal stroke Hospitalisation for heart failure eGFR >2 years Amputation Fracture Acute Kidney Injury
CREDENCE (Perkovic 2019) (n=4401)	Adults with Type 2 Diabetes CKD and eGFR 30-90 and ACR A3	26 weeks	Mean eGFR 56.2 ml/min/1.73 m2 Mean ACR 104.8mg/mmol	Canagliflozin 100mg	Placebo	Renal composite (as above) Cardiovascular composite (as above) All-cause mortality Cardiovascular death Hospitalisation for heart failure End stage kidney disease Doubling serum creatinine Dialysis

Trial [1], author, year, sample size	Population[2]	Follow-up	Baseline eGFR and ACR (mean/median/n%)	Intervention	Comparator	Outcome[2]
						Diabetic ketoacidosis Amputation Fracture Acute Kidney Injury eGFR 6 months
Subgroup of DAPA-CKD (Wheeler 2021) N=4304	Adults with Type 2 Diabetes and CKD with eGFR 25-75	2.4 years (median)	Mean eGFR 43.8 ml/min/1.73 m2 Median ACR 114.64 mg/mmol	Dapagliflozin (10mg)	Placebo	All-cause mortality Cardiovascular death End stage kidney disease eGFR reduction >50% Diabetic ketoacidosis Fracture Hypoglycaemia
Subgroup of DECLARE-TIMI (Wiviott 2019) N=1265	Adults with Type 2 Diabetes and CKD with eGFR <60	4.2 years (median)	Mean eGFR 51.4 ml/min/1.73 m2 ACR not measured at baseline for all patients	Dapagliflozin (10mg)	Placebo	eGFR 6 months Cardiovascular composite (as above) eGFR >2 years

Trial [1], author, year, sample size	Population[2]	Follow-up	Baseline eGFR and ACR (mean/median/n%)	Intervention	Comparator	Outcome[2]
DELIGHT (Pollock 2019) N=293	Adults with Type 2 Diabetes and CKD with eGFR 20-80 or ACR A3	24 weeks	Mean eGFR 49.0 ml/min/1.73 m2 Median ACR 29.8 mg/mmol	Dapagliflozin (10mg)	Placebo	Percentage change from baseline ACR 6 months Diabetic ketoacidosis Amputation Fracture Hypoglycaemia Genitourinary infection
DERIVE (Fioretto 2018) N=321	Adults with Type 2 Diabetes and CKD witheGFR 45-60	24 weeks	Mean eGFR 53.5ml/min/1.73 m2 Median ACR 2.97 mg/mmol	Dapagliflozin 10mg	Placebo	eGFR 6 months Diabetic ketoacidosis Fracture Hypoglycaemia Genitourinary infection
Subgroup of EMPA-REG (Wanner 2018) (N=2250)	Adults with Type 2 Diabetes and CKD with eGFR 30-60 or ACR A1&A2, A3	3.1 years (mean)	Subgroup eGFR 30-60: Mean eGFR 54.4 ml/min/1.73 m2 A1 = 37.7% A2 = 27.35% A3 = 34.3%	Empagliflozin 10mg	Placebo	Cardiovascular composite (as above) All cause mortality Cardiovascular death Hospitalisation for heart failure

Trial [1], author, year, sample size	**Population[2]**	**Follow-up**	**Baseline eGFR and ACR (mean/median/n%)**	**Intervention**	**Comparator**	**Outcome[2]**
						Fatal/non-fatal MI Fatal/non-fatal stroke
VERTIS RENAL (Grunberger 2018) n=467	Adults with Type 2 Diabetes and CKD with eGFR 30-60	1 year	Mean eGFR 46.6 ml/min/1.73 m2 ACR not reported at baseline	Ertugliflozin 5mg and 15mg	Placebo	eGFR 6 months Hypoglycaemia Genitourinary infection
YALE 2013/14 n=269	Adults with Type 2 Diabetes and CKD with eGFR 30-50	26 weeks / 52 weeks	eGFR 39.9 ml/min/1.73 m2 Mean ACR 30.6 mg/mmol	Canagliflozin 100/300 mg	Placebo	eGFR 6 months Genitourinary infection

1. Trial name and primary paper where evidence table is saved under (see appendix D). Multiple papers were used to find relevant outcomes for each trial – see section 1.1.14 for full list.
2. Subgroup populations and outcomes listed are relevant for this review.

1 See appendix D for full evidence tables

1.1.6 Summary of the effectiveness evidence

Outcomes were stratified by eGFR and ACR at baseline when data was available. Stratified results are only presented in the table below when there was evidence for a difference between subgroups (test for subgroup differences p<0.05). Forest plots, including those for all stratifications are shown in Appendix E. Full GRADE profiles are shown in Appendix F.

Table 3: Summary of effectiveness evidence

SGLT2 inhibitor	Placebo	Effect size (95% CI)	Quality	Interpretation of effect[9]
Renal composite – end stage kidney disease, doubling serum creatinine, renal death				
n=1110	n=929	HR 0.71 (0.59-0.85)	Moderate[1]	Favours SGLT2
Cardiovascular composite: 3-point MACE				
n=5040	n=4485	HR 0.81 (0.73-0.91)	Moderate[1]	Favours SGLT2
All-cause mortality				
n=4414	n=4402	HR 0.80 (0.69-0.93)	High	Favours SGLT2
Cardiovascular death				
n=5434	n=5422	HR 0.83 (0.71-0.97)	High	Favours SGLT2
Non-fatal myocardial infarction				
n=2202	n=2199	HR 0.81 (0.59-1.11)	Moderate[2]	Cannot differentiate
Non-fatal stroke				
n=2202	n=2199	HR 0.80 (0.56-1.15)	Moderate[2]	Cannot differentiate
Fatal/non-fatal myocardial infarction[10]				
n=2232	n=1627	HR 0.78 (0.59-1.02)	Moderate[2]	Cannot differentiate
Fatal/non-fatal stroke				
n=2232	n=1627	HR 0.70 (0.49-0.98)	High	Favours SGLT2
Hospitalisation for heart failure				
n=3979	n=3971	HR 0.58 (0.48-0.71)	High	Favours SGLT2
End stage kidney disease				
n=3657	n=3650	HR 0.68 (0.57-0.82)	High	Favours SGLT2
Doubling of serum creatinine				
n=2202	n=2199	HR 0.60 (0.48-0.75)	High	Favours SGLT2
eGFR reduction >50%				
n=1455	n=1451	HR 0.53 (0.42-0.67)	High	Favours SGLT2
Dialysis				
n=3657	n=3653	HR 0.72 (0.57-0.90)	High	Favours SGLT2
eGFR at 6 months (without eGFR stratification)				
n=3126	n=2986	MD -1.91 (-2.83, -0.99)	High	Favours placebo
eGFR at 6 months – eGFR 30-45				
n=228	n=249	MD -2.30 (-3.02, -1.58)	Moderate[7]	Favours placebo
eGFR at 6 months – eGFR 45-60				
n=888	n=761	MD -4 (-4.22, -3.77)	High	Favours placebo

SGLT2 inhibitor	Placebo	Effect size (95% CI)	Quality	Interpretation of effect[9]
eGFR at last available data point >2years				
n=790	n=581	MD 2.37 (1.75-2.98)	Moderate[3]	Favours SGLT2
eGFR at 6 months ACR A2				
n=1232	n=876	MD -1 (-1.71, -0.29)	High	Favours Placebo
eGFR at 6 months ACR A3				
n=373	n=325	MD -3 (-3.70, -2.30)	Moderate[7]	Favours Placebo
eGFR at last available data point >2 years ACR A2				
n=349	n=159	MD 3 (1.89-4.11)	Moderate[7]	Favours SGLT2
eGFR at last available data point >2 years ACR A3				
n=63	N=26	MD 10 (7.86-12.14)	High	Favours SGLT2
Percentage change from baseline ACR 6 months (%)				
n=148	n=144	MD -1.00 (-24.84, 22.84)	High	Cannot differentiate
Percentage change from baseline ACR last available data point >2years (%)				
n=276	n=115	MD -38.00 (-81.24, 5.24)	High	Cannot differentiate
Diabetic ketoacidosis – Canagliflozin				
11 / 2200	1 / 2197	RR 10.98 (1.42 – 85.01)	High	Favours placebo
Diabetic ketoacidosis – Dapagliflozin				
0 / 2455	2 / 2457	RR 0.67 (0.11 – 4.00)	Low[4]	Cannot differentiate
Diabetic ketoacidosis - SGLT2 class (Canagliflozin and Dapagliflozin)				
12/4655	3/4654	RR 2.27 (0.21-24.73)	Very Low [3, 5]	Cannot differentiate
Amputation – Canagliflozin				
104 / 3220	79 / 3217	RR 1.48 (0.70 – 3.13)	Very Low [4, 5]	Cannot differentiate
Amputation – Dapagliflozin				
36/2297	38/2294	RR 0.94 (0.60-1.48)	Low[5]	Cannot differentiate
Amputation – SGLT2 class (Canagliflozin and Dapagliflozin)				
140/5517	117/5511	RR 1.28 (0.81 – 2.02)	Low[3, 6]	Cannot differentiate
Fracture				
194/5674	172/5675	RR 1.13 (0.92 – 1.38)	Moderate[6]	Cannot differentiate
Acute Kidney Injury				
96/3220	113/3217	RR 0.85 (0.65-1.11)	Moderate[6]	Cannot differentiate
Hypoglycaemia				
101/2608	142/2771	RR 0.91 (0.72-1.15)	Low[3, 6]	Cannot differentiate
Genitourinary infection				
356/2908	282/2750	RR 1.18 (1.02–1.37)	Low[3, 6]	Favours placebo

1) composite outcome with differences between outcome components, rated down.
2) 95% confidence interval crosses the MID (the line of no effect = 1), rated down for imprecision

SGLT2 inhibitor	Placebo	Effect size (95% CI)	Quality	Interpretation of effect[9]
3) I^2 between 33.3% and 66.7%, rated down for inconsistency. 4) I^2 above 66.7, rated down twice for inconsistency 5) 95% confidence interval crosses the MID (0.8-1.25) at both ends, rated down twice for imprecision 6) 95% confidence interval crosses the MID (0.8-1.25) at one end, rated down for imprecision 7) 95% confidence interval crosses the MID (2.4). MID = 0.5 of the median standard deviation of the comparison group , rated down for imprecision 8) Stratification was also undertaken by ACR and eGFR on other outcomes not shown - no differences in effect were found. Details of full stratified analysis can be found in appendix E - forest plots 9) Hazard ratios and risk ratios below 1 (with 95% confidence intervals below 1) indicate an effect that favours treatment; mean differences (other than eGFR continuous) below 0 (with 95% confidence intervals below 0) indicate an effect that favours treatment. eGFR continuous outcomes below 0 favour placebo. Where 95% CI's crossover 1 (HR/RR) or 0 (MD) we are unable to differentiate between treatment and placebo. Where the effect and 95% CI's lie above 1 or 0, the effect favours placebo. 10) Some studies report non-fatal MI as a separate outcome; other studies group this together with fatal MI.				

1.1.7 Economic evidence

1.1.7.1 Included studies

A systematic search was performed to identify economic evidence for the review question, with 66 papers identified. Following an initial review of titles and abstracts, 5 papers were selected for screening on full text. Following the full text review, one paper (Willis et al. 2021) was identified as an applicable cost-utility analysis for the review question; details of this study are summarised in section 1.1.8. The study selection is shown in more detail in Appendix G, while full economic evidence tables along with the checklists for study applicability and study limitations are shown in Appendix H.

1.1.7.2 Excluded studies

Four papers (McEwan et al. 2020; McEwan et al. 2021; McEwan et al. 2015; Johnston et al. 2018) were screened by full text before being excluded. In all four studies, the primary reason for their exclusion was due to the analyses being based on a general type 2 diabetes population and not limited to a diabetic population with kidney disease.

1.1.8 Summary of included economic evidence

Table 4: Summary of Willis et al. (2021)

Willis et al. (2021). Cost-Effectiveness of Canagliflozin Added to Standard of Care for Treating Diabetic Kidney Disease (DKD) in Patients with Type 2 Diabetes Mellitus (T2DM) in England: Estimates Using the CREDEM-DKD Model	
Study details	**Analysis:** Cost-utility analysis **Approach to analysis:** A discrete event simulation model with an adjustable time horizon. Simulated patients are characterised to be representative of patients in the CREDENCE trial. Baseline characteristics, patient history and time-varying risk factors are used to determine renal health state. The renal health states are supplemented with additional health states relating to MI, stroke, hospitalisation for heart failure and death. DKD progression is experienced in terms of eGFR decline and ACR increase, **Complications considered:** CKD stages 1, 2, 3A, 3B, 4, and 5 prior to dialysis, receiving dialysis, and post renal transplant. Also has health states relating to MI, stroke, hospitalisation for heart failure and death. **Perspective:** United Kingdom **Time horizon:** 10 years **Discounting:** 3.5%
Interventions	**Intervention 1:** Canagliflozin 100mg + SoC

Willis et al. (2021). Cost-Effectiveness of Canagliflozin Added to Standard of Care for Treating Diabetic Kidney Disease (DKD) in Patients with Type 2 Diabetes Mellitus (T2DM) in England: Estimates Using the CREDEM-DKD Model	
	Intervention 2: Current standard of care (SoC)
Population	**Population:** Adults age 30+ with Type 2 diabetes and CKD, defined as: eGFR: 30 to 90 mL/min per 1.73 m^2 and ACR > 30 mg/mmol **Characteristics:** Mean age: 63; Mean diabetes duration: 15.8; Female: 33.9%
Data sources	**Baseline/natural history:** From CREDENCE population **Risk equations:** Extrapolations of eGFR and log(ACR) were informed by linear mixed model risk equations estimated from CREDENCE, with a minimum threshold of eGFR set at which all patients would immediately assigned to start dialysis. Risk equations were estimated from CREDENCE, and use baseline characteristics, patient history, eGFR and log(ACR) to predict events. **Effectiveness:** From CREDENCE **Costs:** Cardiovascular complications taken from Alva et al. (2015); dialysis and kidney transplant taken from Kerr et al. (2012); CKD stages taken from NICE technology appraisal of tolvaptan for treating autosomal dominant polycystic kidney disease for dialysis and kidney transplant costs [TA358]. **QoL:** Uses a range of sources sourced from a targeted literature search.
Base-case results	
Sensitivity analyses	**Deterministic:** Eight sensitivity analyses performed to check the robustness of the model including varying time horizon, treatment effects on renal, CV, and mortality outcomes, including treatment effects for stroke, dialysis and mortality, removal of eGFR fail-safe floor, and assuming same trajectory of eGFR for both arms. **Probabilistic:** Model structure is patient-level, capturing first and second order uncertainty.
Comments	**Source of funding:** This study was financed by Mundipharma and the fees for the journal's Rapid Service was supported by Napp Pharmaceuticals Limited (part of the Mundipharma Network) **Limitations:** Minor limitations with one potentially serious limitation (see Appendix H)

1.1.9 Economic model

A published decision-analytical model (Willis et al. 2021), originally developed to assess the cost-effectiveness of canagliflozin in diabetic kidney disease (DKD) using data from the CREDENCE trial, was adapted to assess the cost-effectiveness of SGLT2 inhibitors compared with standard care for adults with DKD. The rationale for economic modelling, the methodology adopted, the results and the conclusions from this economic analysis are described in Appendix I. This section provides a summary of the methods employed and the results of the economic analysis.

Overview of methods

A decision-analytic model using a microsimulation approach was adapted to evaluate the cost-effectiveness of SGLT2 inhibitors in addition to standard care over a 10-year time horizon. The population comprised adults with DKD and type 2 diabetes mellitus (T2DM), with baseline characteristics based on the patients enrolled in the CREDENCE trial. In the base case population, patients had estimated glomerular filtration rate (eGFR) between 30 to 90 mL/min/1.73 m^2, and urinary albumin to creatinine ratio (ACR) greater than 30 mg/mmol.

The model captured the patients' progression through DKD health states through the projection of eGFR decline and ACR increase. In DKD stage 5, patients could receive dialysis or transplant. The patients also faced risks of all-cause mortality (ACM), myocardial infarction (MI), stroke and hospitalisation for heart failure (HHF).

Disease progression and event rates for patients on standard care were based on risk equations developed from the CREDENCE trial. The relative impact of SGLT2 inhibitors was modelled using evidence from the guideline systematic literature review where possible, and used evidence on treatment effectiveness of canagliflozin from CREDENCE elsewhere.

The perspective for costs and outcomes was that of the NHS and PSS. The outcome of the analysis was the number of quality-adjusted life-years (QALYs) gained. Utility values for DKD health states and DKD-related cardiovascular events were obtained from the literature. Unit costs and resource use were obtained from national sources, the current NICE guideline for CKD and the published literature.

All analyses consisted of simulating 500 cohorts of 500 hypothetical patients, and results were presented in the form of incremental cost-effectiveness ratios (ICERs). The time horizon for assessing outcomes was ten years and future outcomes were discounted at 3.5%.

A series of scenario analyses explored the impact of alternative assumptions and sources of data for the costs and utility values, the application of a treatment waning effect after four years, and varying the time horizon over which outcomes were assessed. An exploratory subgroup analysis of patients with baseline ACR 3-30 mg/mmol was also undertaken, to assess the possibility of extending the recommendation to these patients.

Findings of the economic analysis

The economic analysis indicated that SGLT2 inhibitors are a dominant treatment option compared with standard care alone, being both cost saving and producing more benefits in an analysis based on evidence for patients with ACR > 30 mg/mmol. An exploratory analysis of patients with ACR 3-30 mg/mmol indicated a possibility for SGLT2 inhibitors to be cost-effective in this subgroup; however, the economic analysis was based on less robust evidence and a firm conclusion could not be made.

Strengths and limitations

The economic analysis estimated SGLT2 inhibitors likely to be cost-effective, and found the results robust to a wide range of assumptions. There was uncertainty regarding the modelled predictions of eGFR decline beyond the duration of the trial from which the risk equations were estimated. However, a sensitivity analysis that limited the relative benefit of SGLT2 inhibitors after four years found that it remained a cost-effective use of resources. Assessment of the cost-effectiveness over different time horizons also supported the conclusions, unless a time horizon was used that was insufficiently long enough to capture the downstream benefits of SGLT2 inhibitors.

Limitations in the clinical evidence informing the subgroup analysis of patients with ACR 3-30 mg/mmol meant that the results of this exploratory analysis were less robust than that of the base case analysis. These limitations included progression of DKD for patients on standard care modelled using data for the population with ACR > 30 mg/mmol, and less certainty in the clinical evidence of benefit.
Full details of the economic model can be found in Appendix I.

1.1.10 The committee's discussion and interpretation of the evidence

1.1.1.10.1. The outcomes that matter most

The committee agreed that the outcomes most important from a clinical perspective were mortality, CKD progression and cardiovascular events. While reduction in proteinuria was also considered important, the committee agreed that this would fall under CKD progression and that it was not a patient important outcome. Adverse event outcomes such as fracture or falls were considered as both clinically and economically significant due to costs incurred. Other adverse events such as infections and diabetic ketoacidosis were also considered clinically important.

The committee noted that renal bone disease may be a theoretical long-term effect of taking SGLT2 inhibitors due to an increase in phosphate levels, but this outcome was considered less important for decision making. Renal calcification and anaemia were considered less important for decision making for this review.

1.1.10.2 The quality of the evidence

The evidence was analysed by SGLT2 inhibitors as a class but where significant heterogeneity was found, SGLT2 inhibitor medication subgroups were reported. It was noted that evidence for sotagliflozin was not reported on due to this being an SGLT2 and SGLT1 inhibitor. No studies reported evidence for the effect in children or different ethnic subgroups. SGLT2 inhibitors are not currently licensed for used in people under 18 years. The committee made a research recommendation for evidence stratified by ethnicity (see section 1.1.12.5 for further discussion).

The committee noted that the overall quality of the evidence for cardiovascular events and renal outcomes ranged from moderate to high, while adverse outcomes were rated low to moderate, with all RCTs found to have a low risk of bias. Outcomes were rated down mainly due to imprecision and the committee also noted that some outcomes were reported in one study only, though some trials had large numbers of participants which increased the committee's confidence in the results. There was heterogeneity for some outcomes that was explored with subgroup analysis. If the heterogeneity could not be explained the outcome was rated down for inconsistency. The 3 Point MACE and Renal composite outcomes were rated down (under 'other considerations') because they are composite outcomes and so combine data across outcomes with different clinical consequences, making them less useful for decision making. The separate outcome components were also reported individually (where available) which would provide more clarity. The committee also noted that there were large differences in mean or median baseline ACR levels across the studies, but the trials that specifically recruited people with CKD were mostly in a population with A3 proteinuria (ACR>30mg/mol).

Overall, the quality of evidence was sufficient to support a recommendation for SGLT2 inhibitors to be offered to people with type 2 diabetes and A3 proteinuria. The committee were less confident in the evidence for people with A2 proteinuria and so recommended that SGLT2 inhibitors should be considered for this group. For further details, see sections 1.1.12.3 and 1.1.12.4.

1.1.10.3 Benefits and harms

Overall, the committee agreed that the clinical evidence favoured SGLT2 inhibitors over placebo for adults with CKD and type 2 diabetes. There was evidence for clinically important benefits in terms of CKD progression (including end stage kidney disease, renal death, dialysis, doubling of serum creatinine and a reduction of eGFR>50%), cardiovascular events (including stroke, hospitalisation for heart failure and cardiovascular death) and all-cause mortality. The evidence could not differentiate the effect of SGLT2 inhibitors on myocardial infarctions or non-fatal strokes (though when non-fatal and fatal strokes were considered together, the effect did favour SGLT2 inhibitors over placebo).

The evidence on eGFR favoured placebo over SGLT2 inhibitors at 6 months following treatment onset, but this effect was reversed in an analysis pooling data at the end of the study periods (provided that they were at least 2 years long), which favoured SGLT2 inhibitors over placebo. The committee noted that the short-term reduction in eGFR with SGLT2 inhibitors was small and unlikely to be clinically important in most cases. They also noted that there was no evidence of increased rates of acute kidney injury associated with SGLT2 inhibitors, however this may have been because the trials were too small to identify an effect, rather than because an effect wasn't present.

The committee discussed the evidence for diabetic ketoacidosis (DKA) and the apparent subgroup effect showing for different medication within the SGLT2 inhibitor class. There was agreement that any effect on DKA was biologically likely to be a class effect, and several other things could have influenced this outcome, including: baseline diabetes treatment such as ACE inhibitors or ARBs, the length of and severity of disease, HbA1c levels, the use of insulin or how sick the population were. The committee therefore agreed that a class level recommendation for SGLT2 inhibitors was appropriate.

The committee were reassured that there appeared to be no increased risk of amputation as the evidence was unable to differentiate between placebo and SGLT2 inhibitors. They also noted a consistent effect for fracture across all trials. Again, this may have been because the trials were too small to identify an effect, rather than because an effect wasn't present. The committee noted that hypoglycaemia in people with CKD was likely to be associated with issues of insulin clearance rather than the effect of SGLT2 inhibitors, which was consistent with the lack of effect in the evidence that was presented. The committee agreed that genitourinary infection was an adverse effect of SGLT2 inhibitors and this matched their experience in practice and the Summary of Product Characteristics (SPC) for these medications. The size of the effect in the evidence presented was small, but the committee noted that larger effects of SGLT2 inhibitors on this outcome have been observed in larger trials of SGLT2 inhibitors in broader population groups.

Subgroup analysis was conducted on eGFR categories 30-45, 45-60, 60-90 ml/min/1.73 m^2 and ACR A1, A2 and A3, however the committee noted that not all studies reported outcomes in this way, lowering their confidence in the effect estimate for these subgroups. Only 1 trial reported outcomes specifically for the A2 subgroup, and no trials reported data for the A1 subgroup in a population with CKD. The DAPA-CKD study recruited participants with an ACR of greater than 22.6 mg/mmol which was between the A2 and A3 categories. However, the median ACR for the group with type 2 diabetes (that met the inclusion criteria for this review) was 116 mg/mmol (IQR 53 to 23), and so most participants in this trial would have fallen into the A3 category.

It was noted that while the evidence of treatment effect in subgroup A3 was clear and of good quality, there was less certainty for the A2 subgroup with the effect estimate crossing the line of no effect on some outcomes. It was noted however that there was no significant heterogeneity between these subgroups for most of the important outcomes. For the outcome eGFR at more than 2 years following treatment onset, there was evidence of a different effect across proteinuria subgroups, with a greater benefit of SGLT2 inhibitors in the A3 group. The committee highlighted that event rates would naturally be lower in A2 and A1 populations and data was not available to demonstrate any renal benefits in the A1 population. It was noted that the benefits of treatment were clear in all eGFR subgroups. The effects of SGLT2 inhibitors on amputation, acute kidney injury and fracture were similar across eGFR subgroups and could not be differentiated from placebo. Data on adverse events was available for the A3 subgroup only, but did not appear to differ substantially from the overall analysis.

Overall, the committee felt the evidence of benefit for the A3 population (people with ACR greater than 30 mg/mmol) outweighed the harms providing it was prescribed responsibly, considering the individual patient circumstances and potential for adverse effects as discussed above. A recommendation that SGLT2 inhibitors should be offered was therefore made. There was less certainty in the clinical evidence of benefit and evidence for cost effectiveness for an A2 population (people with ACR between 3 and 30 mg/mol), therefore a weaker ‘consider’ recommendation was made for this group.

No evidence was available specifically for the A1 subgroup, therefore no practice recommendation was made for this group, but a research recommendation was made, recommending that the effectiveness of SGLT2 inhibitors should be investigated for this group.

It was also highlighted that most people with type 2 diabetes and CKD would have a 10% or greater QRISK score for cardiovascular disease, meaning those in an A2 population would be prescribed an SGLT2 anyway due to an existing draft recommendation in the diabetes guideline. For both A2 and A3 groups, the committee felt that eGFR levels should be monitored in people given SGLT2 inhibitors due to an initial dip in levels at 6 months compared with placebo, demonstrated by the evidence, and in line with the recommendations in the British national formulary for SGLT2 inhibitors.

1.1.10.4 Cost effectiveness and resource use

The economic analysis indicated that SGLT2 inhibitors are likely to be cost-effective, being a dominant treatment option compared with standard care alone, i.e. result in both cost savings and greater numbers of benefits, for patients with ACR greater than 30 mg/mol. The committee felt that there was uncertainty in the extrapolation of clinical outcomes beyond the follow-up period of the trials, as these were not supported by clinical evidence and that they were unsure whether the predictions were plausible. However, they understood that it was important to capture the downstream benefits of the intervention, which had a large impact on the results of the analysis. The cost-effectiveness of SGLT2 inhibitors across a wide range of scenarios, including those which limited the duration of treatment effect and evaluated SGLT2 inhibitors over shorter time frames, provided the committee with reassurance that the results were robust despite these limitations.

A subgroup analysis of patients with ACR between 3 and 30 mg/mol provided some tentative evidence that SGLT2 inhibitors may be cost-effective in this group. However, there were a number of limitations in the clinical evidence underpinning the economic analysis that meant that the committee was less confident that the results of this analysis were reliable. Firstly, the progression of DKD for patients on standard care was modelled using data for the population with ACR > 30 mg/mmol. The committee noted that the analysis of studies in the clinical evidence review suggested that there was a statistically significant difference in eGFR progression between ACR subgroups, therefore, the baseline disease progression may not be representative of these patients. Secondly, there was less certainty in the clinical evidence of benefit in this subgroup, and the committee noted that not all studies reported outcomes for ACR populations in a consistent manner. Thus the committee decided that the cost-effectiveness evidence in this subgroup was associated with substantial uncertainty. Further to this, the committee were aware that a recommendation in this subgroup may have a large potential resource impact due to the population size. The committee referred to the principle outlined in 7.2 in ‘Developing NICE guidelines: the manual’ (2014) and agreed that it would want to be increasingly certain of the cost effective of a technology as the resource impact of adoption increases. On this basis, the committee preferred to make a recommendation to consider the use of SGLT2 inhibitors rather than a stronger recommendation that they should be offered in this subgroup.

1.1.10.5 Other factors the committee took into account

Committee members commented that that some ethnicities were at higher risk of macro/microvascular complications for a given level of eGFR, and so may benefit differently from SGLT2 inhibitors. As there was no specific evidence on different ethnic subgroups the committee felt that a research recommendation for SGLT2 inhibitors in people of different ethnicities should be made. The committee also urged caution in the use of SGLT2 inhibitors for people who are frail and with multi-morbidities. Postural hypotension was considered an issue for this population and although not reported on in the evidence looked at by the committee, this was known to be an adverse effect of SGLT2 inhibitors as highlighted in the SPC and evidence from broader populations.

Most of the participants in the trials that were included in this evidence review were taking ARBs or ACE inhibitors at baseline. The committee agreed that it was best practice to offer an ACE inhibitor or ARB at an optimised dose to people with CKD and type 2 diabetes before prescribing an SGLT2 inhibitor and included this in the recommendation.

1.1.11 Recommendations supported by this evidence review

This evidence review supports recommendations 1.8.12 to 1.8.15 and the research recommendations on SGLT2 inhibitors in people with CKD and type 2 diabetes.

1.1.12 References – included studies

1.1.12.1 Effectiveness evidence

Wanner, Cristoph (2016) Empagliflozin and progression of kidney disease in type 2 diabetes. New england journal of medicine. 375 (4) (pp 323-334), 2016. Date of publication: 28 jul 2016.

Bae, Jae Hyun, Park, Eun-Gee, Kim, Sunhee et al. (2019) Effects of Sodium-Glucose Cotransporter 2 Inhibitors on Renal Outcomes in Patients with Type 2 Diabetes: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. Scientific reports 9(1): 13009

Bakris, George, Oshima, Megumi, Mahaffey, Kenneth W. et al. (2020) Effects of canagliflozin in patients with baseline eGFR <30 ml/min per 1.73 m2: Subgroup analysis of the randomized CREDENCE trial. Clinical Journal of the American Society of Nephrology 15(12): 1705-1714

Cherney, David Z I, Cooper, Mark E, Tikkanen, Ilkka et al. (2018) Pooled analysis of Phase III trials indicate contrasting influences of renal function on blood pressure, body weight, and HbA1c reductions with empagliflozin. Kidney international 93(1): 231-244

Cherney, David Z. I., Charbonnel, Bernard, Cosentino, Francesco et al. (2021) Effects of ertugliflozin on kidney composite outcomes, renal function and albuminuria in patients with type 2 diabetes mellitus: an analysis from the randomised VERTIS CV trial. Diabetologia

Dekkers, Claire C J, Wheeler, David C, Sjostrom, C David et al. (2018) Effects of the sodium-glucose co-transporter 2 inhibitor dapagliflozin in patients with type 2 diabetes and Stages 3b-4 chronic kidney disease. Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association 33(11): 2005-2011

Fioretto, Paola, Del Prato, Stefano, Buse, John B et al. (2018) Efficacy and safety of dapagliflozin in patients with type 2 diabetes and moderate renal impairment (chronic kidney disease stage 3A): The DERIVE Study. Diabetes, obesity & metabolism 20(11): 2532-2540

Fioretto, Paola, Stefansson, Bergur V, Johnsson, Eva et al. (2016) Dapagliflozin reduces albuminuria over 2 years in patients with type 2 diabetes mellitus and renal impairment. Diabetologia 59(9): 2036-9

Grunberger, George, Camp, Sarah, Johnson, Jeremy et al. (2018) Ertugliflozin in Patients with Stage 3 Chronic Kidney Disease and Type 2 Diabetes Mellitus: The VERTIS RENAL Randomized Study. Diabetes Therapy 9(1): 49-66

Jardine, Meg J., Mahaffey, Kenneth W., Neal, Bruce et al. (2017) The canagliflozin and renal endpoints in diabetes with established nephropathy clinical evaluation (CREDENCE) study rationale, design, and baseline characteristics. American Journal of Nephrology 46(6): 462-472

Jardine, MJ, Zhou, Z, Mahaffey, KW et al. (2020) Renal, Cardiovascular, and Safety Outcomes of Canagliflozin by Baseline Kidney Function: a Secondary Analysis of the CREDENCE Randomized Trial. Journal of the American Society of Nephrology : JASN 31(5): 1128-1139

Kadowaki, Takashi, Nangaku, Masaomi, Hantel, Stefan et al. (2019) Empagliflozin and kidney outcomes in Asian patients with type 2 diabetes and established cardiovascular disease: Results from the EMPA-REG OUTCOME R trial. Journal of diabetes investigation 10(3): 760-770

Kelly, Michael S, Lewis, Jelena, Huntsberry, Ashley M et al. (2019) Efficacy and renal outcomes of SGLT2 inhibitors in patients with type 2 diabetes and chronic kidney disease. Postgraduate medicine 131(1): 31-42

Levin, Adeera, Perkovic, Vlado, Wheeler, David C. et al. (2020) Empagliflozin and cardiovascular and kidney outcomes across KDIGO risk categories: Post hoc analysis of a randomized, double-blind, placebo-controlled, multinational trial. Clinical Journal of the American Society of Nephrology 15(10): 1433-1444

Mahaffey, Kenneth W, Jardine, Meg J, Bompoint, Severine et al. (2019) Canagliflozin and Cardiovascular and Renal Outcomes in Type 2 Diabetes Mellitus and Chronic Kidney Disease in Primary and Secondary Cardiovascular Prevention Groups. Circulation 140(9): 739-750

Malik, Aaqib H, Yandrapalli, Srikanth, Goldberg, Michael et al. (2020) Cardiovascular Outcomes With the Use of Sodium-Glucose Cotransporter-2 Inhibitors in Patients With Type 2 Diabetes and Chronic Kidney Disease: An Updated Meta-Analysis of Randomized Controlled Trials. Cardiology in review 28(3): 116-124

Neuen, Brendon L, Ohkuma, Toshiaki, Neal, Bruce et al. (2021) Relative and Absolute Risk Reductions in Cardiovascular and Kidney Outcomes With Canagliflozin Across KDIGO Risk Categories: Findings From the CANVAS Program. American journal of kidney diseases : the official journal of the National Kidney Foundation 77(1): 23-34e1

Neuen, Brendon L, Ohkuma, Toshiaki, Neal, Bruce et al. (2018) Cardiovascular and Renal Outcomes With Canagliflozin According to Baseline Kidney Function. Circulation 138(15): 1537-1550

Neuen, Brendon L, Ohkuma, Toshiaki, Neal, Bruce et al. (2019) Effect of Canagliflozin on Renal and Cardiovascular Outcomes across Different Levels of Albuminuria: Data from the CANVAS Program. Journal of the American Society of Nephrology : JASN 30(11): 2229-2242

Neuen, Brendon L, Young, Tamara, Heerspink, Hiddo J L et al. (2019) SGLT2 inhibitors for the prevention of kidney failure in patients with type 2 diabetes: a systematic review and meta-analysis. The lancet. Diabetes & endocrinology 7(11): 845-854

Oshima, Megumi, Jardine, Meg J., Agarwal, Rajiv et al. (2021) Insights from CREDENCE trial indicate an acute drop in estimated glomerular filtration rate during treatment with canagliflozin with implications for clinical practice. Kidney International 99(4): 999-1009

Oshima, Megumi, Neuen, Brendon L, Li, JingWei et al. (2020) Early Change in Albuminuria with Canagliflozin Predicts Kidney and Cardiovascular Outcomes: A Post Hoc Analysis from the CREDENCE Trial. Journal of the American Society of Nephrology : JASN 31(12): 2925-2936

Perkovic, Vlado, de Zeeuw, Dick, Mahaffey, Kenneth W et al. (2018) Canagliflozin and renal outcomes in type 2 diabetes: results from the CANVAS Program randomised clinical trials. The lancet. Diabetes & endocrinology 6(9): 691-704

Perkovic, Vlado, Jardine, Meg J, Neal, Bruce et al. (2019) Canagliflozin and Renal Outcomes in Type 2 Diabetes and Nephropathy. The New England journal of medicine 380(24): 2295-2306

Pollock, Carol, Stefansson, Bergur, Reyner, Daniel et al. (2019) Albuminuria-lowering effect of dapagliflozin alone and in combination with saxagliptin and effect of dapagliflozin and saxagliptin on glycaemic control in patients with type 2 diabetes and chronic kidney disease (DELIGHT): a randomised, double-blind, placebo-controlled trial. The lancet. Diabetes & endocrinology 7(6): 429-441

Toyama, Tadashi, Neuen, Brendon L, Jun, Min et al. (2019) Effect of SGLT2 inhibitors on cardiovascular, renal and safety outcomes in patients with type 2 diabetes mellitus and chronic kidney disease: A systematic review and meta-analysis. Diabetes, obesity & metabolism 21(5): 1237-1250

Wang, Chen, Zhou, Yue, Kong, Zili et al. (2019) The renoprotective effects of sodium-glucose cotransporter 2 inhibitors versus placebo in patients with type 2 diabetes with or without prevalent kidney disease: A systematic review and meta-analysis. Diabetes, obesity & metabolism 21(4): 1018-1026

Wanner, C, Inzucchi, SE, Zinman, B et al. (2020) Consistent Effects of Empagliflozin on Cardiovascular and Kidney Outcomes Irrespective of Diabetic Kidney Disease Categories - Insights from the EMPA-REG OUTCOME trial. Diabetes, obesity & metabolism

Wanner, Christoph, Lachin, John M, Inzucchi, Silvio E et al. (2018) Empagliflozin and Clinical Outcomes in Patients With Type 2 Diabetes Mellitus, Established Cardiovascular Disease, and Chronic Kidney Disease. Circulation 137(2): 119-129

Wheeler, David C, Stefansson, Bergur V, Jongs, Niels et al. (2021) Effects of dapagliflozin on major adverse kidney and cardiovascular events in patients with diabetic and non-diabetic chronic kidney disease: a prespecified analysis from the DAPA-CKD trial. The lancet. Diabetes & endocrinology 9(1): 22-31

Wiviott, Stephen D, Raz, Itamar, Bonaca, Marc P et al. (2019) Dapagliflozin and Cardiovascular Outcomes in Type 2 Diabetes. The New England journal of medicine 380(4): 347-357

Yale, J-F, Bakris, G, Cariou, B et al. (2014) Efficacy and safety of canagliflozin over 52 weeks in patients with type 2 diabetes mellitus and chronic kidney disease. Diabetes, obesity & metabolism 16(10): 1016-27

Yale, J-F, Bakris, G, Cariou, B et al. (2013) Efficacy and safety of canagliflozin in subjects with type 2 diabetes and chronic kidney disease. Diabetes, obesity & metabolism 15(5): 463-73

Yamada, Takayuki, Wakabayashi, Mako, Bhalla, Abhinav et al. (2021) Cardiovascular and renal outcomes with SGLT-2 inhibitors versus GLP-1 receptor agonists in patients with type 2 diabetes mellitus and chronic kidney disease: a systematic review and network meta-analysis. Cardiovascular Diabetology 20(1): 14

Yamout, Hala, Perkovic, Vlado, Davies, Melanie et al. (2014) Efficacy and safety of canagliflozin in patients with type 2 diabetes and stage 3 nephropathy. American journal of nephrology 40(1): 64-74

Ye, Nan, Jardine, Meg J., Oshima, Megumi et al. (2021) Blood Pressure Effects of Canagliflozin and Clinical Outcomes in Type 2 Diabetes and Chronic Kidney Disease: Insights from the CREDENCE Trial. Circulation

Yu, Baisong, Dong, ChunXia, Hu, ZhiJuan et al. (2021) Effects of sodium-glucose co-transporter 2 (SGLT2) inhibitors on renal outcomes in patients with type 2 diabetes mellitus and chronic kidney disease: A protocol for systematic review and meta-analysis. Medicine 100(8): e24655

1.1.12.2 Economic

Alva ML, Gray A, Mihaylova B, Leal J, Holman RR. The impact of diabetes-related complications on healthcare costs: new results from the UKPDS (UKPDS 84). Diabet Med. 2015 Apr;32(4):459-66.

Johnston R, Uthman O, Cummins E, et al. Corrigendum: Canagliflozin, dapagliflozin and empagliflozin monotherapy for treating type 2 diabetes: systematic review and economic evaluation. Health Technol Assess. 2018;21(2):219-220.

Kerr M, Bray B, Medcalf J, O'Donoghue DJ, Matthews B. Estimating the financial cost of chronic kidney disease to the NHS in England. Nephrol Dial Transplant. 2012;27(3):73–80.

McEwan P, Bennett H, Khunti K, et al. Assessing the cost-effectiveness of sodium–glucose cotransporter-2 inhibitors in type 2 diabetes mellitus: A comprehensive economic evaluation using clinical trial and real-world evidence. Diabetes, Obes Metab. 2020;22(12):2364-2374.

McEwan P, Bennett H, Ward T, Bergenheim K. Refitting of the UKPDS 68 risk equations to contemporary routine clinical practice data in the UK. Pharmacoeconomics. 2015;33(2):149-161.

McEwan P, Morgan AR, Boyce R, et al. The cost-effectiveness of dapagliflozin in treating high-risk patients with type 2 diabetes mellitus: An economic evaluation using data from the DECLARE-TIMI 58 trial. Diabetes, Obes Metab. 2021;23(4):1020-1029.

NICE Technology Appraisal 358: Tolvaptan for treating autosomal dominant polycystic kidney disease. 2015. Available at: https://www.nice.org.uk/guidance/ta358 [Accessed July 2021].

Willis M, Nilsson A, Kellerborg K, et al. Cost-Effectiveness of Canagliflozin Added to Standard of Care for Treating Diabetic Kidney Disease (DKD) in Patients with Type 2 Diabetes Mellitus (T2DM) in England: Estimates Using the CREDEM-DKD Model. Diabetes Ther. 2021;12(1):313-328.

Appendices

Appendix A – Review protocol

Review protocol for SGLT2 inhibitors for people with chronic kidney disease and type 2 diabetes

ID	Field	Content
0.	PROSPERO registration number	This review protocol was not registered on PROSPERO.
1.	Review title	SGLT2 inhibitors for people with chronic kidney disease and type 2 diabetes
2.	Review question	What is the clinical and cost effectiveness of SGLT2 inhibitors for children, young people and adults with chronic kidney disease and type 2 diabetes?
3.	Objective	To determine the effectiveness and cost effectiveness of SGLT2 inhibitors for people with chronic kidney disease and type 2 diabetes
4.	Searches	The following databases will be searched: • Cochrane Central Register of Controlled Trials (CENTRAL)

		• Cochrane Database of Systematic Reviews (CDSR) • Embase • MEDLINE Searches will be restricted by: • English language • Human studies Searches will not be limited by date. The full search strategies for MEDLINE database will be published in the final review.
5.	Condition or domain being studied	Chronic kidney disease and type 2 diabetes.
6.	Population	Adults, children, and young people with CKD and type 2 diabetes.. Exclusion: • people receiving renal replacement therapy (RRT) • people with acute kidney injury combined with rapidly progressive glomerulonephritis • people receiving palliative care.

		• Pregnant women with type 2 diabetes. • People with type 1 diabetes. • People with type 2 diabetes who are hyperglycaemic and require rescue treatment. • Trials including a mixed population (for example of people with and without CKD or with and without type 2 diabetes) will be excluded unless a subgroup analysis for people with CKD and type 2 diabetes is reported, or this population comprises >85% of the trial population.
7.	Intervention/Exposure/Test	• Existing therapy + SGLT2 inhibitors (including canagliflozin, dapagliflozin, ertugliflozin , empagliflozin) Examples of existing therapy could include antidiabetic medications, such as metformin, or therapy to treat CKD, such as an ARB or ACEi
8.	Comparator/Reference standard/Confounding factors	• Existing therapy • Existing therapy plus Placebo
9.	Types of study to be included	• RCTs • Systematic reviews of RCTs
10.	Other exclusion criteria	• Population exclusions – as listed above • Trials including treatments not available (or no longer available) in the UK

		• Abstracts, conference presentations and theses • Study designs not listed above • Non-human studies • Non-English language studies
11.	Context	The update of CG182 chronic kidney disease in adults: assessment and management currently includes review evidence and recommendations on SGLT2 inhibitors for people with proteinuria and type 2 diabetes. This review will expand that to all people with CKD and type 2 diabetes and look at the cost-effectiveness of this intervention.

12.	Primary outcomes (critical outcomes)	• Proteinuria (measured as urinary albumin to creatinine ratio, ACR)
		• CKD progression: occurrence of end stage kidney disease (End stage renal disease or end stage kidney disease as reported by the study – we will report doubling of serum creatinine, renal replacement therapy and transplant separately if data is available)
		• Mortality (all-cause and cardiovascular)
		• Specific morbidity (at the longest timepoint reported by the study):
		  ○ advancement of renal bone disease,
		  ○ vascular calcification,
		  ○ cardiovascular impact, including macrovascular events (non-fatal MI, non-fatal stroke, hospitalisation for heart failure*)*

		- anaemia - health-related quality of life - eGFR (continuous outcome, or dichotomous outcome as number of participants with eGFR reduction>40% or 50%, as reported) - Adverse outcomes: - Acute kidney injury, - drug specific: hypotension/falls, hypoglycaemia, amputations, genitourinary infections, fractures, diabetic ketoacidosis Mortality outcomes, cardiovascular outcomes, CKD progression and reduction in eGFR by 40% or 50% will be reported as hazard ratios over the course of the study. Other dichotomous outcomes will be reported at the longest timepoint reported by the study. In case of heterogeneity, data will be split by timepoint (<2 years, >2 years). The eGFR and ACR continuous outcomes will be reported 6 months after treatment onset and for the longest available time point, providing it is over 2 years.

		Where trials report data for more than one dose, the dose within the BNF recommended range (taking account correction for renal function) will be chosen for inclusion in the analysis. If a trial only reports on doses outside of the recommended range, committee members will be consulted to decide whether the trial should be included in the meta-analysis, or reported separately.
13.	Secondary outcomes (important outcomes)	None
14.	Data extraction (selection and coding)	The full text of potentially eligible studies will be retrieved and will be assessed in line with the criteria outlined above. Data will be extracted from the included studies for assessment of study quality and evidence synthesis. Extracted information will include: study setting; study population and participant demographics and baseline characteristics; study methodology; recruitment and study completion rates; outcomes and times of measurement and information for assessment of the risk of bias. Study investigators may be contacted for missing data where time and resources allow.
15.	Risk of bias (quality) assessment	Risk of bias will be assessed using the Cochrane RoB 2.0 checklist as described in Developing NICE guidelines: the manual.

16.	Strategy for data synthesis	Meta-analyses of outcome data will be conducted for all comparators that are reported by more than one study, with reference to the Cochrane Handbook for Systematic Reviews of Interventions (Higgins et al. 2021). Fixed- and random-effects models (der Simonian and Laird) will be fitted for all comparators, with the presented analysis dependent on the degree of heterogeneity in the assembled evidence. Fixed-effects models will be the preferred choice to report, but in situations where the assumption of a shared mean for fixed-effects model is clearly not met, even after appropriate pre-specified subgroup analyses is conducted, random-effects results are presented. Fixed-effects models are deemed to be inappropriate if one or both of the following conditions was met: • Significant between study heterogeneity in methodology, population, intervention or comparator was identified by the reviewer in advance of data analysis. • The presence of significant statistical heterogeneity in the meta-analysis, defined as $I^2 \geq 50\%$. Meta-analyses will be performed in Cochrane Review Manager V5.3.

17.	Analysis of sub-groups	**Stratifications** Where it is possible to disambiguate data, data will be stratified by: • ethnicity (White, Asian (other than South East Asian), South East Asian, African/Caribbean, mixed or other ethnic background) • different albuminuria categories at baseline: A1 (ACR <30 mg/g) A2 (ACR 30-300 mg/g) and A3 (ACR >300 mg/g) • different eGFR categories at baseline (eGFR 30-45,45-60 and >60 ml/min/1.73 m2) **Subgroup analysis** In the case of unexplained heterogeneity, data will be split by type of SGLT2 inhibitor (the primary analysis will treat SGLT2 inhibitors as a class). See outcomes section on dealing with heterogeneity due to different follow up periods. If heterogeneity cannot be explained, a random effects model will be used. If stratification/subgroup analysis results in very few studies in each category that will not be helpful for decision making, subgroup analysis will not be performed and if heterogeneity remains, a random effects model will be used.
18.	Type and method of review	☒ Intervention ☐ Diagnostic ☐ Prognostic

		☐ Qualitative ☐ Epidemiologic ☐ Service Delivery ☐ Other (please specify)		
19.	Language	English		
20.	Country	England		
21.	Anticipated or actual start date	December 2019		
22.	Anticipated completion date	March 2020		
23.	Stage of review at time of this submission	**Review stage**	**Started**	**Completed**
		Preliminary searches	☐	☐

		Piloting of the study selection process	☐	☐
		Formal screening of search results against eligibility criteria	☐	☐
		Data extraction	☐	☐
		Risk of bias (quality) assessment	☐	☐
		Data analysis	☐	☐
24.	Named contact	**5a. Named contact** NICE Guideline Updates Team **5b Named contact e-mail** GUTprospero@nice.org.uk **5e Organisational affiliation of the review** National Institute for Health and Care Excellence (NICE)		

25.	Review team members	From the Guideline Updates Team: • Dr Kathryn Hopkins • Mr Anthony Gildea • Ms Lucy Beggs • Ms Sarah Glover • Mr Kusal Lokuge
26.	Funding sources/sponsor	This Systematic review is being completed by the Guideline Updates Team which is part of NICE.
27.	Conflicts of interest	All guideline committee members and anyone who has direct input into NICE guidelines (including the evidence review team and expert witnesses) must declare any potential conflicts of interest in line with NICE's code of practice for declaring and dealing with conflicts of interest. Any relevant interests, or changes to interests, will also be declared publicly at the start of each guideline committee meeting. Before each meeting, any potential conflicts of interest will be considered by the guideline committee Chair and a senior member of the development team. Any decisions to exclude a person from all or part of a meeting will be documented. Any changes to a member's declaration of interests will be recorded in the minutes of the meeting. Declarations of interests will be published with the final guideline.
28.	Collaborators	Development of this systematic review will be overseen by an advisory committee who will use the review to inform the development of evidence-based recommendations in line with section 3 of Developing NICE guidelines: the manual. Members of the guideline committee are available on the NICE website.
29.	Other registration details	

30.	Reference/URL for published protocol	
31.	Dissemination plans	NICE may use a range of different methods to raise awareness of the guideline. These include standard approaches such as: • notifying registered stakeholders of publication • publicising the guideline through NICE's newsletter and alerts • issuing a press release or briefing as appropriate, posting news articles on the NICE website, using social media channels, and publicising the guideline within NICE.
32.	Keywords	Chronic kidney disease, Type 2 Diabetes, SGLT2 inhibitors
33.	Details of existing review of same topic by same authors	
34.	Current review status	☒ Ongoing ☐ Completed but not published

		☐ Completed and published ☐ Completed, published and being updated ☐ Discontinued
35..	Additional information	
36.	Details of final publication	www.nice.org.uk

Appendix B – Literature search strategies

1 exp Diabetes Mellitus, Type 2/

2 (Type* adj4 ("2" or "II" or two*) adj4 (diabete* or diabetic*)).tw.

3 ((Maturit* or adult* or slow*) adj4 onset* adj4 (diabete* or diabetic*)).tw.

4 ((Ketosis-resistant* or stable*) adj4 (diabete* or diabetic*)).tw.

5 ((Non-insulin* or Non insulin* or Noninsulin*) adj4 depend* adj4 (diabete* or diabetic*)).tw.

6 NIDDM.tw.

7 or/1-6

8 Sodium-Glucose Transporter 2/

9 (Sodium* adj4 Glucose* adj4 Transporter* adj4 "2").tw.

10 (Sodium* adj4 Glucose* adj4 (co-transporter* or cotransporter* or co transporter*) adj4 "2").tw.

11 (SGLT* or gliflozin*).tw.

12 Canagliflozin/

13 (Canagliflozin* or Invokana* or Dapagliflozin* or Forxiga* or Ertugliflozin* or Steglatro* or Empagliflozin* or Jardiance* or Glyxambi*).tw.

14 or/8-13

15 exp Renal Insufficiency, Chronic/

16 ((chronic* or progressi*) adj1 (renal* or kidney*)).tw.

17 ((kidney* or renal*) adj1 insufficien*).tw.

18 ckd*.tw.

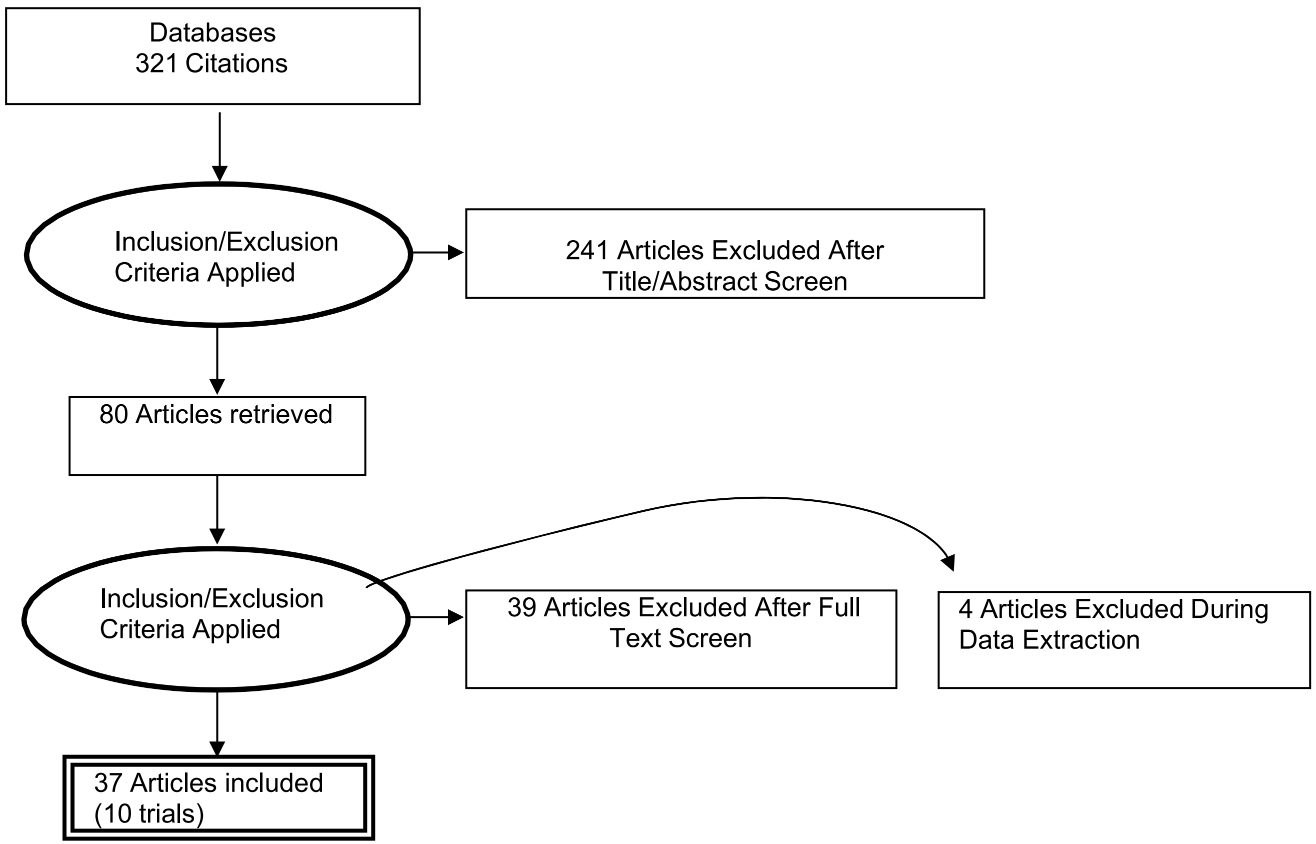
19 ((kidney* or renal*) adj1 fail*).tw.

20 ((endstage* or end-stage* or "end stage*") adj1 (renal* or kidney*)).tw.

21 (esrd* or eskd*).tw.

22	"Chronic Kidney Disease-Mineral and Bone Disorder"/
23	or/15-22
24	7 and 14 and 23
25	animals/ not humans/
26	24 not 25
27	limit 26 to english language
28	randomized controlled trial.pt.
29	randomi?ed.mp.
30	placebo.mp.
31	or/28-30
32	(MEDLINE or pubmed).tw.
33	systematic review.tw.
34	systematic review.pt.
35	meta-analysis.pt.
36	intervention$.ti.
37	or/32-36
38	31 or 37
39	27 and 38

Appendix C – Effectiveness evidence study selection



Appendix D – Evidence tables and risk of bias

Cherney, 2021

Bibliographic Reference	Cherney, David Z. I.; Charbonnel, Bernard; Cosentino, Francesco; Dagogo-Jack, Samuel; McGuire, Darren K.; Pratley, Richard; Shih, Weichung J.; Frederich, Robert; Maldonado, Mario; Pong, Annpey; Cannon, Christopher P.; Effects of ertugliflozin on kidney composite outcomes, renal function and albuminuria in patients with type 2 diabetes mellitus: an analysis from the randomised VERTIS CV trial; Diabetologia; 2021

Study details

Other publications associated with this study included in review	Grunberger, George; Camp, Sarah; Johnson, Jeremy; Huyck, Susan; Golm, Gregory; Engel, Samuel S.; Lauring, Brett; Terra, Steven G.; Mancuso, James P.; Jiang, Zhi Wei; Ertugliflozin in Patients with Stage 3 Chronic Kidney Disease and Type 2 Diabetes Mellitus: The VERTIS RENAL Randomized Study; Diabetes Therapy; 2018; vol. 9 (no. 1); 49-66
Trial registration number and/or trial name	VERTIS CV Reg no NCT01986881
Study type	Randomised controlled trial (RCT)
Study location	Region of enrollment, n (%) North America 1812 (22.0) Europe (including Russia) 4632 (56.2) Asia 522 (6.3) Australia/New Zealand 173 (2.1) South and Central America 722 (8.8) South Africa 377 (4.6) - 34 countries
Study setting	567 centres (no further details)
Study dates	Enrolment - Dec 2013 - July 2015 and June 2016 - April 2017 The final follow-up window was from September 2019 through December 2019; the last patient visit took place on December 27, 2019

Sources of funding	Merck Sharp & Dohme and Pfizer
Inclusion criteria	Patients ≥40 years of age at the time of the initial Screening visit (V1) with a diagnosis of type 2 diabetes mellitus (T2DM) in accordance with American Diabetes Association (ADA) guidelines. Diabetes, Glycated hemoglobin (HbA1c) at the Screening visit (V1) of 7.0–10.5% (53–91 mmol/mol) on stable allowable antihyperglycemic agent(s) (AHA) or on no background AHA for at least 8 weeks prior to the Screening visit (V1). eGFR ≥30 ml min−1 [1.73 m]−2. Atherosclerosis - evidence or a history of atherosclerosis involving the coronary, cerebral, or peripheral vascular systems as follows (must have at least 1 of the following): a. Coronary artery disease as indicated by a history of presumed spontaneous myocardial infarction (MI; hospitalized with final diagnosis of MI, excluding peri-procedural or definite secondary MI [eg, due to profound anemia or hypertensive emergency, troponin increase in sepsis] in which the most recent event occurred at least 3 months (90 days) prior to the Screening visit (V1); OR b. Coronary artery disease as indicated by a history of coronary revascularization through either a percutaneous coronary intervention at least 3 months (90 days) prior to the Screening visit (V1) or coronary artery bypass graft at least 3 months (90 days) prior to the Screening visit (V1); OR c. Ischemic (presumed thrombotic) cerebrovascular disease as indicated by a history of ischemic stroke (hospitalized with a final diagnosis of nonhemorrhagic stroke [includes completion of a standard evaluation for stroke in an acute care facility or stroke clinic without hospital admission] with the most recent event occurring at least 3 months (90 days) prior to the Screening visit (V1) or a history of carotid revascularization at least 3 months (90 days) prior to the Screening visit (V1); OR d. Peripheral arterial disease as indicated by: 1. Angiographically-documented peripheral vascular disease; or 2. Resting ankle/brachial index of b0.85 (measured by a certified vascular laboratory) plus symptoms of claudication; or 3. Amputation, peripheral bypass, or peripheral angioplasty of the extremities secondary to ischemia occurring at least 3 months (90 days) prior to the Screening visit (V1).

Exclusion criteria	eGFR - estimated glomerular filtration rate below 30 ml per minute per 1.73 m2 of body-surface area. Diabetes - history of type 1 diabetes or ketoacidosis
Intervention(s)	Eligible patients were randomly assigned in a 1:1:1 ratio to receive 5 mg or 15 mg of ertugliflozin, added to background standard-of-care treatment
Comparator	matching placebo once daily, added to background standard-of-care treatment
Outcome measures	decline in eGFR Composite kidney outcome doubling of baseline serum creatinine, kidney dialysis/transplant or renal death eGFR changes in eGFR over time Albuminuria changes over time Mortality All cause and CV Heart failure End-stage Kidney Disease Doubling of serum creatinine Urine Albumin Creatinine Ratio HbA1c Change from Baseline in HbA1c at Week 18, Week 52 and annually thereafter.

	Proportion of patients with HbA1c <7% (53 mmol/mol) and <6.5% (48 mmol/mol) at 12, 24 and 36 months and annually thereafter (not presented in this report). Hospitalisation Blood pressure Body weight Composite CV outcome - death from cardiovascular causes, nonfatal myocardial infarction, or nonfatal stroke (i.e., a major adverse cardiovascular event). (MACE) Time to glycemic rescue therapy during first 18 weeks of study time to initiation of insulin and change in insulin dose from baseline at Week 18, Week 52 and annually thereafter
Number of participants	8246
Duration of follow-up	3 years (median), 3.5 (mean)
Loss to follow-up	Of the randomised individuals, 87–88% completed the trial alive, 8.5–9.2%died and 3.7–4.1% withdrew. Study medication was discontinued prematurely in 27.9% and 23.5% of participants in the placebo and ertugliflozin groups, respectively.
Methods of analysis	Intention to treat The time-to-event endpoints were analysed using a stratified Cox proportional hazards model, including treatment group as a covariate with cohort as a stratification factor (cohort one [participants randomised before protocol amendment, between December 2013 and July 2015] and cohort two [participants randomised after protocol amendment, in 2016 and beyond]).

Study arms

Ertugliflozin 5mg (N = 2752)

Study dates	Enrolment December 2013 through July 2015 and from June 2016 through April 2017 The final follow-up window was from September 2019 through December 2019	Enrolment December 2013 through July 2015 and from June 2016 through April 2017 The final follow-up window was from September 2019 through December 2019	Enrolment December 2013 through July 2015 and from June 2016 through April 2017 The final follow-up window was from September 2019 through December 2019

5mg

Placebo (N = 2747)

Duration of follow-up	3 years (median)	3 years (median)	3 years (median)

Ertugliflozin (N = 2747)

15mg

Ertugliflozin pooled (N = 5499)

Characteristics

Study-level characteristics

Characteristic	Study (N = 8246)
% Female Sample size	n = 2477 ; % = 30
Mean age (SD) Mean (SD)	64.4 (8.1)
BMI Mean (SD)	32 (5.4)
White Sample size	n = 7232 ; % = 87.8
Black Sample size	n = 235 ; % = 2.9
Asian Sample size	n = 497 ; % = 6
Other Sample size	n = 274 ; % = 3.3
Hispanic or Latino Sample size	n = 1042 ; % = 12.6

Arm-level characteristics

Characteristic	Ertugliflozin 5mg (N = 2752)	Placebo (N = 2747)	Ertugliflozin 15mg (N = 2747)	Ertugliflozin pooled (N = 5499)
CKD stage 3 Sample size	*empty data*	n = 608; % = 22.1	*empty data*	n = 1199 ; % = 21.8
Age - CKD stage 3 Mean (SD)	*empty data*	68 (7.5)	*empty data*	68.3 (7.6)
Female - CKD stage 3 Sample size	*empty data*	n = 211 ; % = 34.7	*empty data*	n = 440 ; % = 36.7
eGFR (MDRD) - CKD 3 Mean (SD) (ml/min/1.73m2)	*empty data*	48.6 (8)	*empty data*	49.1 (8)
UACR - CKD stage 3 (mg/mmol) Median (IQR)	*empty data* (*empty data* to *empty data*)	3.5 (0.9 to 13.2)	*empty data* (*empty data* to *empty data*)	3.4 (0.9 to 15.4)
BMI (CKD stage 3) (kg/m²) Mean (SD)	*empty data*	32.7 (5.7)	*empty data*	32.3 (5.5)

Critical appraisal - GUT Cochrane Risk of Bias tool (RoB 2.0) Normal RCT

Section	Question	Answer
Domain 1: Bias arising from the randomisation process	1. 1. Was the allocation sequence random?	Yes
Domain 1: Bias arising from the randomisation process	1. 2. Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Probably yes
Domain 1: Bias arising from the randomisation process	1.3 Did baseline differences between intervention groups suggest a problem with the randomisation process?	No
Domain 1: Bias arising from the randomisation process	Risk of bias judgement for the randomisation process	Low
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.1. Were participants aware of their assigned intervention during the trial?	Probably no
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	Probably no
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	Not applicable
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.4. If Y/PY to 2.3: Were these deviations from intended intervention balanced between groups?	Not applicable
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.5 If N/PN/NI to 2.4: Were these deviations likely to have affected the outcome?	Not applicable

Section	Question	Answer
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Yes
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	*N/A.*
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	Risk of bias for deviations from the intended interventions (effect of assignment to intervention)	Low
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.1. Were participants aware of their assigned intervention during the trial?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.3. If Y/PY/NI to 2.1 or 2.2: Were important co-interventions balanced across intervention groups?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.4. Could failures in implementing the intervention have affected the outcome?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.5. Did study participants adhere to the assigned intervention regimen?	*N/A.*

Section	Question	Answer
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.6. If N/PN/NI to 2.3 or 2.5 or Y/PY/NI to 2.4: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	Risk of bias judgement for deviations from the intended interventions (effect of adhering to intervention)	*N/A.*
Domain 3. Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomised?	Yes
Domain 3. Bias due to missing outcome data	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	Not applicable
Domain 3. Bias due to missing outcome data	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Not applicable
Domain 3. Bias due to missing outcome data	3.4 If Y/PY/NI to 3.3: Do the proportions of missing outcome data differ between intervention groups?	Not applicable
Domain 3. Bias due to missing outcome data	3.5 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	Not applicable
Domain 3. Bias due to missing outcome data	Risk-of-bias judgement for missing outcome data	Low
Domain 4. Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	No
Domain 4. Bias in measurement of the outcome	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups ?	Probably no

Section	Question	Answer
Domain 4. Bias in measurement of the outcome	4.3 If N/PN/NI to 4.1 and 4.2: Were outcome assessors aware of the intervention received by study participants ?	*N/A.*
Domain 4. Bias in measurement of the outcome	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	*N/A.*
Domain 4. Bias in measurement of the outcome	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	*N/A.*
Domain 4. Bias in measurement of the outcome	Risk-of-bias judgement for measurement of the outcome	Low
Domain 5. Bias in selection of the reported result	5.1 Was the trial analysed in accordance with a pre-specified plan that was finalised before unblinded outcome data were available for analysis ?	Yes
Domain 5. Bias in selection of the reported result	5.2 Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	No/Probably no
Domain 5. Bias in selection of the reported result	5.3 Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple analyses of the data?	No/Probably no
Domain 5. Bias in selection of the reported result	Risk-of-bias judgement for selection of the reported result	Low
Overall bias and Directness	Risk of bias judgement	Low
Overall bias and Directness	Overall Directness	Directly applicable

Fioretto, 2018

Bibliographic Reference	Fioretto, Paola; Del Prato, Stefano; Buse, John B; Goldenberg, Ronald; Giorgino, Francesco; Reyner, Daniel; Langkilde, Anna Maria; Sjostrom, C David; Sartipy, Peter; DERIVE Study, Investigators; Efficacy and safety of dapagliflozin in patients with type 2 diabetes and moderate renal impairment (chronic kidney disease stage 3A): The DERIVE Study.; Diabetes, obesity & metabolism; 2018; vol. 20 (no. 11); 2532-2540

Study details

Trial registration number and/or trial name	DERIVE Study - NCT02413398
Study type	Randomised controlled trial (RCT)
Study location	8 countries (USA, Canada, Bulgaria, the Czech Republic, Italy, Poland, Spain and Sweden.)
Study setting	88 Centres
Study dates	June 15, 2015 to November 7, 2017
Sources of funding	AstraZeneca
Inclusion criteria	Age ≥18 to <75 years Diabetes - inadequate glycaemic control (HbA1c ≥7.0% and ≤11.0% at screening) Treatment - undergoing a stable glucose-lowering treatment regimen (stable diet and exercise alone or in combination with any approved oral glucose-lowering medication, except SGLT2-inhibitors, and/or long/intermediate-acting insulin or mixed insulin), BMI - body mass index (BMI) of 18–45 kg/m2

	CKD - CKD 3A (eGFR, 40–65 mL/min/1.73 m2 at Visit 1 to enter the lead-in period and eGFR, 45–59 mL/min/1.73 m2 at Visits 1, 2 or 3 to be randomized).
Exclusion criteria	Treatment - The use of metformin was restricted to doses for moderate renal impairment (eGFR, 30–59 mL/min/1.73 m2) according to local guidelines or the investigator's judgement. Patients were excluded if they had received treatment with an SGLT2 inhibitor, a glucagon-like peptide 1 (GLP-1) receptor agonist or a rapid/short-acting insulin at screening. Renal - Certain renal diseases (rapid worsening of renal function from Visit 1 to Visit 3, intercurrent kidney disease other than diabetic nephropathy, renal transplant, dialysis or ultrafiltration). Cardiovascular - history of severe uncontrolled hypertension, certain CV/vascular diseases within 3 months prior to enrolment (myocardial infarction, cardiac surgery or revascularization, unstable angina, unstable heart failure, heart failure Class IV according to the New York Heart Association [NYHA], transient ischaemic attack or significant cerebrovascular disease, unstable or previously undiagnosed arrhythmia), biochemistry and bloods - patients who had a serum potassium level of >5.5 mmol/L, a serum calcium level of <1.99 mmol/L or > ULN, or a haemoglobin level of ≤90 g/L were excluded
Intervention(s)	Patients with T2D were randomized to dapagliflozin 10 mg once daily or matching placebo, taken orally in the morning, in addition to their usual care. Randomization was stratified by pre-enrolment glucose lowering therapy (long/intermediate-acting and mixed insulin regimen, metformin, sulphonylurea, thiazolidinedione or other regimen). Oral glucose-lowering drugs (apart from SGLT2 inhibitors), insulin (apart from rapid/short-acting insulins), antihypertensive drugs, lipid-lowering drugs and anti-platelet drugs were permitted as long as the dose remained constant throughout the 24-week treatment period. Patients who developed a loss of glycaemic control during the 24-week treatment period, defined as fasting plasma glucose (FPG) of >13.3 mmol/L during Weeks 4–12 or FPG of >11.1 mmol/L during Weeks 12–24, were eligible for open-label rescue medication in addition to the blinded treatment. Rescue medication could comprise any appropriate glucose lowering agent, with the exception of SGLT2 inhibitors.
Outcome measures	Adverse events - Safety objectives included adverse events (AEs), serious AEs and AEs of interest, based on a predefined list of preferred terms from the Medical Dictionary for Regulatory Activities (MedDRA), including genital and urinary tract infections, volume depletion, renal impairment/failure, bone fractures and diabetic ketoacidosis (DKA); mean change from baseline in heart rate at 24 weeks; mean change in eGFR from baseline to Week 24 and at the 3-weeks post-treatment follow-up period; the proportion of patients discontinuing study medication because of worsening renal insufficiency, defined as confirmed eGFR of 30 mL/min/1.73 m2, over 24 weeks.

	The proportion of patients who experienced hypoglycaemia events and the frequency of such events were also evaluated. Major hypoglycaemia was defined as a symptomatic episode requiring external assistance because of severe impairment in consciousness or behaviour, with a capillary or plasma glucose value <3.0 mmol/L and prompt recovery after glucose or glucagon administration. Minor hypoglycaemia was defined as either a symptomatic episode with a capillary or plasma glucose value <3.5 mmol/L or a capillary or plasma glucose value <3.5 mmol/L, without symptoms, that does not qualify as a major episode. Other episodes of hypoglycaemia were defined as an episode reported by an investigator that did not meet the criteria for a major or minor episode. HbA1c - The primary efficacy outcome was mean change from baseline in HbA1c at Week 24. Exploratory endpoints included the proportion of patients achieving HbA1c <7% at 24 weeks, change from baseline in urine albumin: creatinine ratio (UACR) at Week 24 (all patients and according to albuminuria status)
Number of participants	321
Duration of follow-up	24 weeks
Loss to follow-up	Most patients completed the study, regardless of discontinuation of doubleblind treatment (156 patients [97.5%] in the dapagliflozin group and 154 patients [95.7%] in the placebo group) and most also completed the 24-week double-blind treatment period (149 patients [93.1%] in the dapagliflozin group and 146 patients [90.7%] in the placebo group).
Methods of analysis	Efficacy analyses were performed on the full analysis set, comprising all randomized patients who received at least 1 dose of double-blind study medication and for whom a baseline value and at least 1 postbaseline efficacy value were available. The primary efficacy analysis, change from baseline in HbA1c at Week 24, was based on a mixed effects model with repeated measures (MMRM) using “direct likelihood” which assumed that missing data were missing at random.

Study arms

Dapagliflozin 10 mg (N = 160)

Placebo (N = 161)

Characteristics

Arm-level characteristics

Characteristic	Dapagliflozin 10 mg (N = 160)	Placebo (N = 161)
Age (years) Mean (SD)	65.3 (*empty data*)	66.2 (*empty data*)
Female Sample size	n = 69 ; % = 43.1	n = 70 ; % = 43.5
White Sample size	n = 141 ; % = 88.1	n = 140 ; % = 87
Black/African-American Sample size	n = 11 ; % = 6.9	n = 12 ; % = 7.5
Asian Sample size	n = 5 ; % = 3.1	n = 8 ; % = 5
American Indian/Alaskan Native Sample size	n = 2 ; % = 1.3	n = 0 ; % = 0

Characteristic	Dapagliflozin 10 mg (N = 160)	Placebo (N = 161)
Other Sample size	n = 1 ; % = 0.6	n = 1 ; % = 0.6
Hispanic or Latino Sample size	n = 33 ; % = 20.6	n = 44 ; % = 27.3
Not hispanic or latino Sample size	n = 127 ; % = 79.4	n = 117 ; % = 72.7
BMI (kg/m^2) Mean (SD)	32.6 (4.7)	31.6 (5)
eGFR (ml/min/1.73m2 of body surface area) Mean (SD)	53.3 (8.7)	53.6 (10.6)
UACR Median (IQR) mg/mmol	23.5 (2.7 to 5852)	29 (3.8 to 8474)
HbA1c (%) Mean (SD)	8.33 (1.08)	8.03 (1.08)

Critical appraisal - GUT Cochrane Risk of Bias tool (RoB 2.0) Normal RCT

Section	Question	Answer
Domain 1: Bias arising from the randomisation process	1. 1. Was the allocation sequence random?	Yes
Domain 1: Bias arising from the randomisation process	1. 2. Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Yes
Domain 1: Bias arising from the randomisation process	1.3 Did baseline differences between intervention groups suggest a problem with the randomisation process?	No
Domain 1: Bias arising from the randomisation process	Risk of bias judgement for the randomisation process	Low
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.1. Were participants aware of their assigned intervention during the trial?	No
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	No
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	No/Probably no

Section	Question	Answer
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.4. If Y/PY to 2.3: Were these deviations from intended intervention balanced between groups?	Not applicable
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.5 If N/PN/NI to 2.4: Were these deviations likely to have affected the outcome?	Not applicable
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Yes
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	Not applicable
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	Risk of bias for deviations from the intended interventions (effect of assignment to intervention)	Low *(intention to treat. Most patients completed the study, regardless of discontinuation of doubleblind treatment (156 patients [97.5%] in the dapagliflozin group and 154 patients [95.7%] in the placebo group) and most also completed the 24-week double-blind treatment period (149 patients [93.1%] in the dapagliflozin group and 146 patients [90.7%] in the placebo group).)*
Domain 2b: Risk of bias due to deviations from the	2.1. Were participants aware of their assigned intervention during the trial?	*N/A.*

Section	Question	Answer
intended interventions (effect of adhering to intervention)		
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.3. If Y/PY/NI to 2.1 or 2.2: Were important co-interventions balanced across intervention groups?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.4. Could failures in implementing the intervention have affected the outcome?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.5. Did study participants adhere to the assigned intervention regimen?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.6. If N/PN/NI to 2.3 or 2.5 or Y/PY/NI to 2.4: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	Risk of bias judgement for deviations from the intended interventions (effect of adhering to intervention)	*N/A.*

Section	Question	Answer
Domain 3. Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomised?	Yes
Domain 3. Bias due to missing outcome data	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	Not applicable
Domain 3. Bias due to missing outcome data	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Not applicable
Domain 3. Bias due to missing outcome data	3.4 If Y/PY/NI to 3.3: Do the proportions of missing outcome data differ between intervention groups?	Not applicable
Domain 3. Bias due to missing outcome data	3.5 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	Not applicable
Domain 3. Bias due to missing outcome data	Risk-of-bias judgement for missing outcome data	Low *(Most patients completed the study, regardless of discontinuation of doubleblind treatment (156 patients [97.5%] in the dapagliflozin group and154 patients [95.7%] in the placebo group) and most also completed the 24-week double-blind treatment period (149 patients [93.1%] in the dapagliflozin group and 146 patients [90.7%] in the placebo group).)*
Domain 4. Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	No

Section	Question	Answer
Domain 4. Bias in measurement of the outcome	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups ?	No
Domain 4. Bias in measurement of the outcome	4.3 If N/PN/NI to 4.1 and 4.2: Were outcome assessors aware of the intervention received by study participants ?	No
Domain 4. Bias in measurement of the outcome	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Not applicable
Domain 4. Bias in measurement of the outcome	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	Not applicable
Domain 4. Bias in measurement of the outcome	Risk-of-bias judgement for measurement of the outcome	Low
Domain 5. Bias in selection of the reported result	5.1 Was the trial analysed in accordance with a pre-specified plan that was finalised before unblinded outcome data were available for analysis ?	Yes
Domain 5. Bias in selection of the reported result	5.2 Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	No/Probably no

Section	Question	Answer
Domain 5. Bias in selection of the reported result	5.3 Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple analyses of the data?	No/Probably no
Domain 5. Bias in selection of the reported result	Risk-of-bias judgement for selection of the reported result	Low
Overall bias and Directness	Risk of bias judgement	Low
Overall bias and Directness	Overall Directness	Directly applicable

Grunberger, 2018

Bibliographic Reference	Grunberger, George; Camp, Sarah; Johnson, Jeremy; Huyck, Susan; Golm, Gregory; Engel, Samuel S.; Lauring, Brett; Terra, Steven G.; Mancuso, James P.; Jiang, Zhi Wei; Ertugliflozin in Patients with Stage 3 Chronic Kidney Disease and Type 2 Diabetes Mellitus: The VERTIS RENAL Randomized Study; Diabetes Therapy; 2018; vol. 9 (no. 1); 49-66

Study details

Other publications associated with this study included in review	Grunberger G, Camp S, Johnson J, Huyck S, Terra SG, Mancuso JP, Jiang ZW, Golm G, Engel SS, Lauring B. Ertugliflozin in patients with stage 3 chronic kidney disease and type 2 diabetes mellitus: the VERTIS RENAL randomized study. Diabetes Therapy. 2018 Feb;9(1):49-66.

	Cherney DZ, Charbonnel B, Cosentino F, Dagogo-Jack S, McGuire DK, Pratley R, Shih WJ, Frederich R, Maldonado M, Pong A, Cannon CP. Effects of ertugliflozin on kidney composite outcomes, renal function and albuminuria in patients with type 2 diabetes mellitus: an analysis from the randomised VERTIS CV trial. Diabetologia. 2021 Jun;64(6):1256-67. Trial registration number and/or trial name
Trial registration number and/or trial name	ClinicalTrials.gov number, NCT01986881 – VERTIS renal
Study type	Randomised controlled trial (RCT)
Study location	34 countries
Study setting	567 centres (no further details reported)
Study dates	December 2013 through July 2015 and from June 2016 through April 2017; The final follow-up window was from September 2019 through December 2019; the last patient visit took place on December 27, 2019.
Sources of funding	Merck Sharp & Dohme and Pfizer
Inclusion criteria	Age - Adults (aged 25 years or older) Diabetes type 2 diabetes Treatment on diet/exercise with or without AHA monotherapy or combination therapy using other AHAs including insulin and sulfonylureas. Patients on metformin at the screening visit were eligible to participate in the trial if their A1C was C 6.5%

	and B 10.0%; however, they were required to undergo a C 10-week metformin wash-off, and they remained eligible if their A1C was C 7.0% and B 10.5% at the end of this period. Cardiovascular with a glycated haemoglobin level of 7.0 to 10.5% and established atherosclerotic cardiovascular disease involving the coronary, cerebrovascular, or peripheral arterial systems. CKD stage 3 CKD (eGFR C 30 and\60 mL/min/1.73m2 calculated using the MDRD equation) with stable renal function. Stable renal function was defined as a change in eGFR\10 mL/min/1.73m2 between screening and visit 3 (week – 2), with eGFR measurement C 30 to\60 mL/min/1.73 m2 at both visits.
Exclusion criteria	Other conditions history of type 1 diabetes mellitus, history of ketoacidosis, renal-related medical history (including nephrotic range proteinuria ([3000 mg/day) with hypoalbuminemia and edema, rapidly progressive glomerulonephritis, lupus nephritis, renal or systemic vasculitis, renal artery stenosis with renovascular hypertension, or ischemic nephropathy, familial renal glucosuria, renal dialysis, renal transplant, or renal disease requiring treatment with immunosuppressive agents), active obstructive uropathy, or an indwelling urinary catheter. Treatment The only prohibited background AHAs were metformin, rosiglitazone, and other SGLT2 inhibitors. eGFR Estimated glomerular filtration rate below 30 ml per minute per 1.73 m2 of body-surface area Diabetes History of type 1 diabetes or ketoacidosis
Intervention(s)	Prior to randomization, eligible patients entered a 2-week single-blind placebo run-in period. Patients with adequate compliance (C 80% based on pill count) were randomized 1:1:1 to ertugliflozin 5 mg, ertugliflozin 15mg, or placebo while continuing a diet/exercise regimen

	and background AHA therapy (if applicable); all blinded study treatments were taken once daily. Following completion of the initial 26-week treatment period, patients entered a 26-week placebo-controlled extension treatment period (phase B, where they continued with their assigned randomized treatment from phase A); the aim of phase B was to gather additional data on the safety and longer-term efficacy of ertugliflozin.
Comparator	Matching placebo once daily, added to background standard-of-care treatment
Outcome measures	eGFR Analysis of the post-treatment eGFR change from baseline was performed in patients in the overall cohort who were on study medication at week 52 and had eGFR results at baseline, week 52 and week 54. Adverse events Safety endpoints included adverse events (AE), including pre-specified AEs and collections of AEs of special interest [symptomatic hypoglycemia and AEs associated with genital mycotic infection (GMI) (gender-specific), urinary tract infection, and hypovolemia]. In addition to symptomatic hypoglycemia, episodes of documented hypoglycemia, defined as episodes with a glucose level B 70 mg/dL with or without symptoms, were also recorded. Pre-defined limits of change (PDLC; criteria based on normal ranges and abnormalities considered clinically meaningful) for pre-specified laboratory and electrocardiogram (ECG) parameters, as well as changes over time in laboratory parameters [including eGFR, low-density lipoprotein cholesterol (LDL-C) and high-density lipoprotein cholesterol (HDL-C)], ECG measurements, and vital signs were assessed. Albuminuria Normal albuminuria, microalbuminuria, and macroalbuminuria were defined as UACR<30, ≥ 30 and ≥ 300, and UACR>300, respectively.

	Mortality Urine Albumin Creatinine Ratio Renal function was further evaluated through urinary albumin/creatinine ratio (UACR) at week 26. HbA1c change from baseline in A1C at week 26 in the overall cohort.
Number of participants	In total - 8250 Underwent randomization; 8246 Were included in the intention-to-treat population.
Duration of follow-up	1 year (main endpoint at week 26)
Loss to follow-up	In total sample - 13% (n=358/2747) in the placebo arm, 12% (n=330/2752) in the ertugliflozin, 5 mg/day are and 12.6% (n=346/2747) ertugliflozin, 15 mg/day arm did not complete the study. ITT analysis undertaken.
Methods of analysis	Intention to treat Stratified Cox proportional-hazards model that included the trial group as a covariate and cohort of enrolment as the stratification factor was used to evaluate the primary outcome. The Kaplan–Meier method was used to estimate the cumulative incidence (first occurrence) of an outcome event over time in each trial group.
Additional comments	

Study arms

Ertugliflozin (5 mg) (N = 158)

Ertugliflozin (15 mg) (N = 155)

Placebo (N = 154)

Characteristics

Study-level characteristics

Characteristic	Study (N = 8246)
Mean age (SD) Mean (SD)	67.3 (8.6)
BMI (kg/m²) Mean (SD)	32.5 (6.1)
Male Sample size	n = 231 ; % = 49.5
White Sample size	n = 380 ; % = 81.4
Asian Sample size	n = 45 ; % = 9.6

Characteristic	Study (N = 8246)
Black or African American Sample size	n = 19 ; % = 4.1
American Indian or Alaska Native Sample size	n = 1 ; % = 0.2
Multiple Sample size	n = 22 ; % = 4.7
Hispanic or Latino Sample size	n = 87 ; % = 18.6

Arm-level characteristics

Characteristic	Ertugliflozin (5 mg) (N = 158)	Ertugliflozin (15 mg) (N = 155)	Placebo (N = 154)
Age Mean (SD)	66.7 (8.3)	67.5 (8.5)	67.5 (8.9)
Male Sample size	n = 84 ; % = 53.2	n = 75 ; % = 48.4	n = 72 ; % = 46.8
White Sample size	n = 127 ; % = 80.4	n = 119 ; % = 76.8	n = 134 ; % = 87

Characteristic	Ertugliflozin (5 mg) (N = 158)	Ertugliflozin (15 mg) (N = 155)	Placebo (N = 154)
Asian Sample size	n = 16 ; % = 10.1	n = 20 ; % = 12.9	n = 9 ; % = 5.8
Black or African American Sample size	n = 6 ; % = 3.8	n = 9 ; % = 5.8	n = 4 ; % = 2.6
American Indian or Alaska Native Sample size	n = 0 ; % = 0	n = 0 ; % = 0	n = 1 ; % = 0.6
Multiple Sample size	n = 9 ; % = 5.7	n = 7 ; % = 4.5	n = 6 ; % = 3.9
Hispanic or Latino Sample size	n = 29 ; % = 18.4	n = 31 ; % = 20	n = 27 ; % = 17.5
BMI (kg/m²) Mean (SD)	32.6 (6.8)	31.7 (5.3)	33.2 (6.1)
eGFR ml/min/1.73m2 Mean (SD)	46.8 (7.8)	46.9 (9.1)	46 (9.4)

Critical appraisal - GUT Cochrane Risk of Bias tool (RoB 2.0) Normal RCT

Section	Question	Answer
Domain 1: Bias arising from the randomisation process	1. 1. Was the allocation sequence random?	Yes
Domain 1: Bias arising from the randomisation process	1. 2. Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Yes
Domain 1: Bias arising from the randomisation process	1.3 Did baseline differences between intervention groups suggest a problem with the randomisation process?	No
Domain 1: Bias arising from the randomisation process	Risk of bias judgement for the randomisation process	Low *((Multicentre, double-blind, randomized, placebo-controlled, event-driven, noninferiority trial; Randomization was performed at a central location with the use of an interactive voice-response system and was based on a computer-generated schedule with randomly permuted blocks, stratified according to geographic region; Study described the baseline characteristics of the patients as well balanced between the ertugliflozin group and the placebo group; However the use of diuretics, were used more often in the placebo group than in the ertugliflozin group at the end of the trial but this is not considered a to be a source of bias; The method of analysis is not specified.))*
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.1. Were participants aware of their assigned intervention during the trial?	No

Section	Question	Answer
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	No
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	No/Probably no
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.4. If Y/PY to 2.3: Were these deviations from intended intervention balanced between groups?	Not applicable
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.5 If N/PN/NI to 2.4: Were these deviations likely to have affected the outcome?	Not applicable
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Yes
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	Not applicable

Section	Question	Answer
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	Risk of bias for deviations from the intended interventions (effect of assignment to intervention)	Low *((Multicentre, double-blind, randomized, placebo-controlled, event-driven, noninferiority trial; Randomization was performed at a central location with the use of an interactive voice-response system and was based on a computer-generated schedule with randomly permuted blocks, stratified according to geographic region; Intention to treat analysis undertaken that considered 99.9% of randomized participants (n=4 participants were excluded post randomization due to being enrolled twice; involved in another ertugliflozin trial)))*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.1. Were participants aware of their assigned intervention during the trial?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.3. If Y/PY/NI to 2.1 or 2.2: Were important co-interventions balanced across intervention groups?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.4. Could failures in implementing the intervention have affected the outcome?	*N/A.*

Section	Question	Answer
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.5. Did study participants adhere to the assigned intervention regimen?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.6. If N/PN/NI to 2.3 or 2.5 or Y/PY/NI to 2.4: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	Risk of bias judgement for deviations from the intended interventions (effect of adhering to intervention)	*N/A.*
Domain 3. Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomised?	Yes
Domain 3. Bias due to missing outcome data	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	Not applicable
Domain 3. Bias due to missing outcome data	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Not applicable
Domain 3. Bias due to missing outcome data	3.4 If Y/PY/NI to 3.3: Do the proportions of missing outcome data differ between intervention groups?	Not applicable

Section	Question	Answer
Domain 3. Bias due to missing outcome data	3.5 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	Not applicable
Domain 3. Bias due to missing outcome data	Risk-of-bias judgement for missing outcome data	Low *(Intention to treat analysis undertaken that considered 99.9% of those randomized; The non-inferiority analysis for the primary outcome considered participants who at received at least one dose of treatment/placebo (99.9%); 12.5% (n=1034) participant did not complete the trial)*
Domain 4. Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	No
Domain 4. Bias in measurement of the outcome	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups ?	Probably no
Domain 4. Bias in measurement of the outcome	4.3 If N/PN/NI to 4.1 and 4.2: Were outcome assessors aware of the intervention received by study participants ?	No
Domain 4. Bias in measurement of the outcome	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Not applicable
Domain 4. Bias in measurement of the outcome	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	Not applicable

Section	Question	Answer
Domain 4. Bias in measurement of the outcome	Risk-of-bias judgement for measurement of the outcome	Low *(Clinical event rates were used to measure all predefined and prespecified outcomes, with all the primary and secondary outcome events centrally adjudicated on by a cardiovascular adjudication committee in a blinded manner; The study is outlined as a multicentre, double-blind, randomized, placebo-controlled, event-driven, noninferiority trial. Randomization was performed at a central location with the use of an interactive voice-response system and was based on a computer-generated schedule with randomly permuted blocks, stratified according to geographic region.)*
Domain 5. Bias in selection of the reported result	5.1 Was the trial analysed in accordance with a pre-specified plan that was finalised before unblinded outcome data were available for analysis ?	Yes
Domain 5. Bias in selection of the reported result	5.2 Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	No/Probably no
Domain 5. Bias in selection of the reported result	5.3 Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple analyses of the data?	No/Probably no
Domain 5. Bias in selection of the reported result	Risk-of-bias judgement for selection of the reported result	Low *((Pre-specified analysis plan is outlined in the paper and published in Cannon et al 2018, with the analysis undertaken is in line with this plan. Primary and secondary outcomes were all prespecified with*

Section	Question	Answer
		outcome measures assessed via clinical event rates with all the primary and secondary outcome events centrally adjudicated on by a cardiovascular adjudication committee in a blinded manner.))
Overall bias and Directness	Risk of bias judgement	Low
Overall bias and Directness	Overall Directness	Directly applicable

Neuen, 2019

Bibliographic Reference	Neuen, Brendon L; Ohkuma, Toshiaki; Neal, Bruce; Matthews, David R; de Zeeuw, Dick; Mahaffey, Kenneth W; Fulcher, Greg; Li, Qiang; Jardine, Meg; Oh, Richard; Heerspink, Hiddo L; Perkovic, Vlado; Effect of Canagliflozin on Renal and Cardiovascular Outcomes across Different Levels of Albuminuria: Data from the CANVAS Program.; Journal of the American Society of Nephrology : JASN; 2019; vol. 30 (no. 11); 2229-2242

Study details

Secondary publication of another included study- see primary study for details	
Other publications associated with this study included in review	Neuen BL, Ohkuma T, Neal B, Matthews DR, de Zeeuw D, Mahaffey KW, Fulcher G, Blais J, Li Q, Jardine MJ, Perkovic V. Relative and absolute risk reductions in cardiovascular and kidney outcomes with canagliflozin across KDIGO risk categories: findings from the CANVAS Program. American Journal of Kidney Diseases. 2021 Jan;77(1):23-34.

	Neuen BL, Ohkuma T, Neal B, Matthews DR, De Zeeuw D, Mahaffey KW, Fulcher G, Desai M, Li Q, Deng H, Rosenthal N. Cardiovascular and renal outcomes with canagliflozin according to baseline kidney function: data from the CANVAS Program. Circulation. 2018 Oct 9;138(15):1537-50. Perkovic V, de Zeeuw D, Mahaffey KW, Fulcher G, Erondu N, Shaw W, Barrett TD, Weidner-Wells M, Deng H, Matthews DR, Neal B. Canagliflozin and renal outcomes in type 2 diabetes: results from the CANVAS Program randomised clinical trials. The lancet Diabetes & endocrinology. 2018 Sep 1;6(9):691-704.
Trial registration number and/or trial name	CANVAS and CANVAS-R. Study numbers NCT01032629 and NCT01989754.
Study type	Randomised controlled trial (RCT)
Study location	30 countries
Study setting	667 centres
Study dates	December 9, 2009 - February 22, 2017
Sources of funding	Janssen Research & Development, LLC
Inclusion criteria	CKD mean eGFR mL/min/1.73m² at baseline: 74.4 (SD 21.3) for people with microalbuminuria and 66.4 (SD 22.3) for people with macroalbuminuria Albuminuria Subgroups with microalbuminuria (urinary albumin/creatinine ratio 30 to <300 mg/g) and macroalbuminuria (urinary albumin/creatinine ratio ≥300 mg/g)

	Age 30 years or older with established atherosclerotic vascular disease, or 50 years or older with 2 or more cardiovascular risk factors (duration of diabetes of at least 10 years, systolic blood pressure higher than 140 mmHg while receiving one or more antihypertensive agents, microalbuminuria or macroalbuminuria, current smoking, or high-density lipoprotein cholesterol level of less than 1 mmol/L) Diabetes Type 2 diabetes Other HbA1c levels ≥7.0% and ≤10.5%
Exclusion criteria	Other conditions eGFR <30 mL/min/1.73 m²
Outcome measures	Composite kidney outcome (1) sustained 40% decline in eGFR, kidney failure, or death due to kidney disease and (2) sustained 40% decline in eGFR, kidney failure, or death due to cardiovascular or kidney disease (i.e., a composite cardiorenal outcome similar to the primary outcome in CREDENCE) 3) a composite of sustained doubling of serum creatinine (sent for adjudication if sustained for two consecutive measures ≥30 days apart or if occurring on the last available measurement), end-stage kidney disease (defined as the composite of maintenance dialysis that was sustained for at least 30 days, renal transplantation, or eGFR <15 mL/minper 1·73 m² sustained for at least 30 days), and death from renal causes (defined as participant death with a proximate renal cause) 4) the composite of each of these outcomes combined with either death from cardiovascular causes or new-onset macroalbuminuria. For each composite outcome, time to the first event was counted, with any subsequent events disregarded. Each of the components of the composite outcomes are also separately reported. eGFR

	Continuous kidney outcome, eGFR slope, defined as the annual mean difference in eGFR between canagliflozin and placebo during acute and chronic treatment periods. eGFR was calculated using the Modification of Diet in Renal Disease (MDRD) Study equation. End points of 40% reduction in eGFR and doubling of serum creatinine were sent for adjudication if sustained for 2 consecutive measures of ≥30 days apart or occurring on the last available measure. Creatinine level Serum creatinine level collected at study visits was centrally measured. End points of 40% reduction in eGFR and doubling of serum creatinine were sent for adjudication if sustained for 2 consecutive measures of ≥30 days apart or occurring on the last available measure. Adverse events Authors reported all serious adverse events for the CANVAS Program along with serious or non-serious adverse events for the CANVAS trial alone due to differences in adverse event reporting between the trials. Renal-related serious adverse events were recorded throughout both trials, and all adverse events (irrespective of seriousness) were also collected in CANVAS until Jan 7, 2014. Renal-related safety outcomes included any (serious and non-serious) renal adverse events (collected from CANVAS until Jan 7, 2014), or serious adverse events and adverse events that led to study drug discontinuation (collected throughout both trials), including acute kidney injury, and were evaluated on the basis of incidence of preferred term, by use of a standard narrow Medical Dictionary for Regulatory Activities (MedDRA) query. Hyperkalaemia was evaluated with the MedDRA preferred terms hyperkalaemia and increased blood potassium. Albuminuria Albuminuria was measured in first-morning void urine specimens and calculated as a UACR. Adverse events, both serious and nonserious, were collected and reported for the CANVAS trial until January 2014, as mandated by the US Food and Drug Administration and other regulatory bodies as a requirement for initial approval for the use of canagliflozin. After this time, only serious adverse events, adverse events leading to study drug discontinuation, or selected adverse events of interest were collected in the CANVAS trial. Albuminuria was measured in first morning void urine specimens and calculated as the UACR. Changes in albuminuria with canagliflozin treatment were calculated as the ratio of the geometric mean of post-randomisation UACR measures compared with the placebo group. New-onset albuminuria was defined as the development of microalbuminuria or macroalbuminuria in participants with baseline normoalbuminuria (defined as UACR of less than 30 mg/g). New-onset microalbuminuria was defined as the development of a UACR of 30–300 mg/g in participants with baseline normoalbuminuria and in whom the UACR had increased by at least 30% from baseline. New-

	onset macroalbuminuria was defined as the development of a UACR greater than 300 mg/g in participants with baseline normoalbuminuria or microalbuminuria and in whom the UACR increased by at least 30% from baseline.
Number of participants	3026
Duration of follow-up	A mean follow-up duration of 188.2 weeks - 3.6 years. The mean follow-up was 188 weeks (SD 106; 296 weeks [SD 74] in CANVAS and 108 weeks [20] in CANVAS-R).
Loss to follow-up	
Methods of analysis	Intention-to-treat
Additional comments	

Study arms

Canagliflozin (N = 5794)

Duration of follow-up	a mean follow-up duration of 188.2 weeks - 3.6 years

100 to 300 mg daily

Placebo (N = 4346)

Duration of follow-up	a mean follow-up duration of 188.2 weeks - 3.6 years

Matching placebo

Characteristics

Study-level characteristics

Characteristic	Study (N =10,140)
eGFR <60 mL/min per 1·73 m² Sample size	n = 2039 ; % = 20.1
eGFR ≥ 60 mL/min per 1·73 m² Sample size	n = 8101 ; % =79.9
eGFR <60 mL/min per 1·73 m² Mean (SD)	49.1 (8)
eGFR ≥ 60 mL/min per 1·73 m² Mean (SD)	83.4 (16.6)
White eGFR <60 mL/min per 1·73 m² Sample size	n = 1673 ; % = 82
Asian eGFR <60 mL/min per 1·73 m² Sample size	n = 216 ; % = 11
Black or African-American eGFR <60 mL/min per 1·73 m² Sample size	n = 46 ; % = 2

Characteristic	Study (N =10,140)
Other eGFR <60 mL/min per 1·73 m² Sample size	n = 104 ; % = 5
White eGFR ≥60 mL/min per 1·73 m² Sample size	n = 6269 ; % = 77
Asian eGFR ≥ 60 mL/min per 1·73 m² Sample size	n = 1068 ; % = 13
Black or African-American eGFR ≥ 60 mL/min per 1·73 m² Sample size	n = 290 ; % = 4
Other eGFR ≥ 60 mL/min per 1·73 m² Sample size	n = 474 ; % = 6
eGFR <60 mL/min per 1·73 m² mg/g Median (IQR)	21.6 (7.7 to 117.8)
eGFR ≥ 60 mL/min per 1·73 m² mg/g Median (IQR)	11.3 (6.5 to 33)
eGFR <60 mL/min per 1·73 m² Sample size	n = 1129 ; % = 56

Characteristic	Study (N =10,140)
eGFR ≥ 60 mL/min per 1·73 m² Sample size	n = 5876 ; % = 73
eGFR <60 mL/min per 1·73 m² Sample size	n = 887 ; % = 44
eGFR ≥ 60 mL/min per 1·73 m² Sample size	n = 2139 ; % = 27

Arm-level characteristics

Characteristic	Canagliflozin (N = 5794)	Placebo (N = 4346)
eGFR (mL/min per 1·73 m²) Mean (SD)	49.1 (8)	83.4 (16.6)
eGFR <60 mL/min per 1·73 m² Mean (SD)	49.2 (7.8)	49 (8.3)
eGFR ≥ 60 mL/min per 1·73 m² Mean (SD)	83.2 (16.5)	83.6 (16.7)
eGFR <60 mL/min per 1·73 m² Sample size	n = 489 ; % = 44	n = 398 ; % = 43

Characteristic	Canagliflozin (N = 5794)	Placebo (N = 4346)
eGFR ≥ 60 mL/min per 1·73 m² Sample size	n = 1239 ; % = 27	n = 900 ; % = 27
eGFR ≥ 60 mL/min per 1·73 m² Mean (SD)	62.1 (8)	62.3 (8)
eGFR <60 mL/min per 1·73 m² Sample size	n = 451 ; % = 41	n = 402 ; % = 43
eGFR ≥ 60 mL/min per 1·73 m² Sample size	n = 1584 ; % = 34	n = 1195 ; % = 35
White eGFR <60 mL/min per 1·73 m² Sample size	n = 907 ; % = 82	n = 766 ; % = 82
Asian eGFR <60 mL/min per 1·73 m² Sample size	n = 118 ; % = 11	n = 98 ; % = 11
Black or African-American eGFR <60 mL/min per 1·73 m² Sample size	n = 27 ; % = 2	n = 19 ; % = 2
Other eGFR <60 mL/min per 1·73 m² Sample size	n = 58 ; % = 5	n = 46 ; % = 5

Characteristic	Canagliflozin (N = 5794)	Placebo (N = 4346)
White eGFR ≥ 60 mL/min per 1·73 m² Sample size	n = 3600 ; % = 77	n = 2669 ; % = 78
Asian eGFR ≥ 60 mL/min per 1·73 m² Sample size	n = 659 ; % = 14	n = 409 ; % = 12
Black or African-American eGFR ≥ 60 mL/min per 1·73 m² Sample size	n = 149 ; % = 3	n = 141 ; % = 4
Other eGFR ≥ 60 mL/min per 1·73 m² Sample size	n = 276 ; % = 6	n = 198 ; % = 6
Normoalbuminuria eGFR <60 mL/min per 1·73 m² Sample size	n = 610 ; % = 56	n = 519 ; % = 57
Normoalbuminuria eGFR ≥ 60 mL/min per 1·73 m² Sample size	n = 3401 ; % = 73	n = 2475 ; % = 73
Microalbuminuria and Macroalbuminuria eGFR <60 mL/min per 1·73 m² Sample size	n = 489 ; % = 44	n = 398 ; % = 43
Microalbuminuria and Macroalbuminuria eGFR ≥ 60 mL/min per 1·73 m² Sample size	n = 1239 ; % = 27	n = 900 ; % = 27

Critical appraisal - GUT Cochrane Risk of Bias tool (RoB 2.0) Normal RCT

Section	Question	Answer
Domain 1: Bias arising from the randomisation process	1. 1. Was the allocation sequence random?	Yes
Domain 1: Bias arising from the randomisation process	1. 2. Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Yes
Domain 1: Bias arising from the randomisation process	1.3 Did baseline differences between intervention groups suggest a problem with the randomisation process?	No
Domain 1: Bias arising from the randomisation process	Risk of bias judgement for the randomisation process	Low
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.1. Were participants aware of their assigned intervention during the trial?	No
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	No
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	Not applicable
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.4. If Y/PY to 2.3: Were these deviations from intended intervention balanced between groups?	Not applicable

Section	Question	Answer
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.5 If N/PN/NI to 2.4: Were these deviations likely to have affected the outcome?	Not applicable
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Yes
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	Not applicable
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	Risk of bias for deviations from the intended interventions (effect of assignment to intervention)	Low
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.1. Were participants aware of their assigned intervention during the trial?	No
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	No
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.3. If Y/PY/NI to 2.1 or 2.2: Were important co-interventions balanced across intervention groups?	Not applicable
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.4. Could failures in implementing the intervention have affected the outcome?	Probably no

Section	Question	Answer
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.5. Did study participants adhere to the assigned intervention regimen?	Probably yes
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.6. If N/PN/NI to 2.3 or 2.5 or Y/PY/NI to 2.4: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	Yes
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	Risk of bias judgement for deviations from the intended interventions (effect of adhering to intervention)	Low
Domain 3. Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomised?	Yes
Domain 3. Bias due to missing outcome data	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	Not applicable
Domain 3. Bias due to missing outcome data	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Not applicable
Domain 3. Bias due to missing outcome data	3.4 If Y/PY/NI to 3.3: Do the proportions of missing outcome data differ between intervention groups?	Not applicable
Domain 3. Bias due to missing outcome data	3.5 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	Not applicable
Domain 3. Bias due to missing outcome data	Risk-of-bias judgement for missing outcome data	Low
Domain 4. Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	No

Section	Question	Answer
Domain 4. Bias in measurement of the outcome	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups ?	No
Domain 4. Bias in measurement of the outcome	4.3 If N/PN/NI to 4.1 and 4.2: Were outcome assessors aware of the intervention received by study participants ?	No
Domain 4. Bias in measurement of the outcome	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Not applicable
Domain 4. Bias in measurement of the outcome	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	Not applicable
Domain 4. Bias in measurement of the outcome	Risk-of-bias judgement for measurement of the outcome	Low
Domain 5. Bias in selection of the reported result	5.1 Was the trial analysed in accordance with a pre-specified plan that was finalised before unblinded outcome data were available for analysis ?	Yes
Domain 5. Bias in selection of the reported result	5.2 Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	No/Probably no
Domain 5. Bias in selection of the reported result	5.3 Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple analyses of the data?	No/Probably no
Domain 5. Bias in selection of the reported result	Risk-of-bias judgement for selection of the reported result	Low
Overall bias and Directness	Risk of bias judgement	Low

Section	Question	Answer
Overall bias and Directness	Overall Directness	Directly applicable

Perkovic, 2019

Bibliographic Reference	Perkovic, Vlado; Jardine, Meg J; Neal, Bruce; Bompoint, Severine; Heerspink, Hiddo J L; Charytan, David M; Edwards, Robert; Agarwal, Rajiv; Bakris, George; Bull, Scott; Cannon, Christopher P; Capuano, George; Chu, Pei-Ling; de Zeeuw, Dick; Greene, Tom; Levin, Adeera; Pollock, Carol; Wheeler, David C; Yavin, Yshai; Zhang, Hong; Zinman, Bernard; Meininger, Gary; Brenner, Barry M; Mahaffey, Kenneth W; CREDENCE Trial, Investigators; Canagliflozin and Renal Outcomes in Type 2 Diabetes and Nephropathy.; The New England journal of medicine; 2019; vol. 380 (no. 24); 2295-2306

Study details

Secondary publication of another included study- see primary study for details	
Other publications associated with this study included in review	Bakris G, Oshima M, Mahaffey KW, Agarwal R, Cannon CP, Capuano G, Charytan DM, de Zeeuw D, Edwards R, Greene T, Heerspink HJ. Effects of Canagliflozin in Patients with Baseline eGFR< 30 ml/min per 1.73 m 2. Jardine MJ, Mahaffey KW, Neal B, Agarwal R, Bakris GL, Brenner BM, Bull S, Cannon CP, Charytan DM, De Zeeuw D, Edwards R. The Canagliflozin and Renal Endpoints in Diabetes with Established Nephropathy Clinical Evaluation (CREDENCE) study rationale, design, and baseline characteristics. American journal of nephrology. 2017;46(6):462-72.

	Jardine MJ, Zhou Z, Mahaffey KW, Oshima M, Agarwal R, Bakris G, Bajaj HS, Bull S, Cannon CP, Charytan DM, de Zeeuw D. Renal, cardiovascular, and safety outcomes of canagliflozin by baseline kidney function: a secondary analysis of the CREDENCE randomized trial. Journal of the American Society of Nephrology. 2020 May 1;31(5):1128-39. Mahaffey KW, Jardine MJ, Bompoint S, Cannon CP, Neal B, Heerspink HJ, Charytan DM, Edwards R, Agarwal R, Bakris G, Bull S. Canagliflozin and cardiovascular and renal outcomes in type 2 diabetes mellitus and chronic kidney disease in primary and secondary cardiovascular prevention groups: results from the randomized CREDENCE trial. Circulation. 2019 Aug 27;140(9):739-50. Oshima M, Jardine MJ, Agarwal R, Bakris G, Cannon CP, Charytan DM, de Zeeuw D, Edwards R, Greene T, Levin A, Lim SK. Insights from CREDENCE trial indicate an acute drop in estimated glomerular filtration rate during treatment with canagliflozin with implications for clinical practice. Kidney international. 2021 Apr 1;99(4):999-1009. Oshima M, Neuen BL, Li J, Perkovic V, Charytan DM, de Zeeuw D, Edwards R, Greene T, Levin A, Mahaffey KW, De Nicola L. Early Change in Albuminuria with Canagliflozin Predicts Kidney and Cardiovascular Outcomes: A Post Hoc Analysis from the CREDENCE Trial. Journal of the American Society of Nephrology. 2020 Dec 1;31(12):2925-36. Perkovic V, Jardine MJ, Neal B, Bompoint S, Heerspink HJ, Charytan DM, Edwards R, Agarwal R, Bakris G, Bull S, Cannon CP. Canagliflozin and renal outcomes in type 2 diabetes and nephropathy. New England Journal of Medicine. 2019 Jun 13;380(24):2295-306. Ye N, Jardine MJ, Oshima M, Hockham C, Heerspink HJ, Agarwal R, Bakris G, Schutte AE, Arnott C, Chang TI, Górriz JL. Blood pressure effects of canagliflozin and clinical outcomes in type 2 diabetes and chronic kidney disease: insights from the CREDENCE trial. Circulation. 2021 May 4;143(18):1735-49.

Trial registration number and/or trial name	CREDENCE (ClinicalTrials.gov Identifier: NCT02065791)
Study type	Randomised controlled trial (RCT)
Study location	34 countries
Study setting	690 sites
Study dates	From March 2014 through May 2017
Sources of funding	Janssen Research and Development
Inclusion criteria	CKD defined as an eGFR 30 to <90 ml per minute per 1.73 m² Albuminuria Urinary albumin/creatinine ratio >300 to 5000 mg/g Age ≥30 years Diabetes Type 2 diabetes Other glycated hemoglobin level of 6.5 to 12.0% (6.5 to 10.5% in Germany, according to a country amendment) Treatment

	stable dose of an angiotensin-converting–enzyme inhibitor or angiotensin-receptor blocker for at least 4 weeks before randomization; a stable dose was considered to be either the maximum labeled dose or a dose not associated with unacceptable side effects
Exclusion criteria	Other conditions Suspected nondiabetic kidney disease or type 1 diabetes, had been treated with immunosuppression for kidney disease, or had a history of dialysis or kidney transplantation Treatment Dual-agent treatment with an angiotensin- converting–enzyme inhibitor and an angiotensin- receptor blocker, a direct renin inhibitor, or a mineralocorticoid-receptor antagonist
Intervention(s)	The use of other background therapy for glycemic management and control of cardiovascular risk factors was recommended in accordance with local guidelines
Comparator	
Outcome measures	Composite kidney outcome Primary endpoint - 1) Composite of ESKD, doubling of serum creatinine, and renal or cardiovascular death Secondary endpoints - 2) the composite of kidney failure, doubling of serum creatinine, or kidney death 3) composite of dialysis, kidney transplantation, or kidney death 3) Composite endpoint of ESKD and renal or cardiovascular death eGFR Change in eGFR over time

	Post hoc analysis: change from baseline in eGFR. In addition, eGFR change was assessed and measured as the acute change in eGFR from baseline to Week 3,6 the annualized chronic change in eGFR from Week 3 until the end of treatment, and the total annualized change in eGFR from baseline to Week 130. Adverse events All adverse events collected and coded using the MedDRA from randomization until 30 days after the last date of blinded study medication. AE of interest - All malignancies, fatal pancreatitis, hemorrhagic/necrotising pancreatitis, severe hypersensitivity reactions (e.g., angioedema, anaphylaxis, Stevens-Johnson syndrome), photosensitivity reactions, serious adverse events of hepatic injury, nephrotoxicity/acute kidney injury, venous thromboembolic events, fractures, diabetic ketoacidosis (and related adverse events including ketoacidosis, metabolic acidosis, or acidosis), amputation, and pregnancy. All episodes of hypoglycemia (both symptomatic and asymptomatic) are recorded on a dedicated hypoglycemia eCRF. Albuminuria Change in albuminuria over time Composite cardiovascular outcome 1) Cardiovascular composite endpoint of cardiovascular death, nonfatal MI, nonfatal stroke, hospitalized congestive heart failure, and hospitalized unstable angina 2) the composite of cardiovascular death or hospitalization for heart failure 3) the composite of cardiovascular death, myocardial infarction, or stroke Mortality 1) cardiovascular death 2) all-cause death 3) Renal death Heart failure hospitalization for heart failure Kidney failure kidney failure; End-stage Kidney Disease

	Doubling of serum creatinine doubling of the serum creatinine level from baseline (average of randomization and pre-randomization value) sustained for at least 30 days according to central laboratory assessment, or kidney death. Myocardial Infarction 1) Fatal and nonfatal MI 2) Hospitalized unstable angina Stroke Fatal and nonfatal stroke
Number of participants	4401
Duration of follow-up	26 weeks
Loss to follow-up	
Methods of analysis	Intention to treat
Additional comments	

Study arms

Canagliflozin (N = 2202)

100 mg orally once daily

Placebo (N = 2199)

Matching placebo

Characteristics

Study-level characteristics

Characteristic	Study (N = 4401)
% Female Sample size	n = 1494 ; % = 33.9
Mean age (SD) Mean (SD)	63 (9.2)
BMI Mean (SD)	31.3 (6.2)
White Sample size	n = 2931 ; % = 66.6
Black Sample size	n = 224 ; % = 5.1
Asian	n = 877 ; % = 19.9

Characteristic	Study (N = 4401)
Sample size	
Other Sample size	n = 369 ; % = 8.4
Estimated GFR ml/min/1.73 m2 Mean (SD)	56.2 (18.2)
Median urinary albumin-to-creatinine ratio Median (IQR)	927 (463 to 1833)

Arm-level characteristics

Characteristic	Canagliflozin (N = 2202)	Placebo (N = 2199)
Age Mean (SD)	62.9 (9.2)	63.2 (9.2)
Female Sample size	n = 762 ; % = 34.6	n = 732 ; % = 33.3
White Sample size	n = 1487 ; % = 67.5	n = 1444 ; % = 65.7
Black	n = 112 ; % = 5.1	n = 112 ; % = 5.1

Characteristic	Canagliflozin (N = 2202)	Placebo (N = 2199)
Sample size		
Asian Sample size	n = 425 ; % = 19.3	n = 452 ; % = 20.6
Other Sample size	n = 178 ; % = 8.1	n = 191 ; % = 8.7
BMI Mean (SD)	31.4 (6.2)	31.3 (6.2)
Estimated GFR — ml/min/1.73 m2 Mean (SD)	56.3 (18.2)	56 (18.3)
Median urinary albumin-to-creatinine ratio (IQR) (mg/g) Median (IQR)	923 (459 to 1794)	931 (473 to 1868)

Critical appraisal - GUT Cochrane Risk of Bias tool (RoB 2.0) Normal RCT

Section	Question	Answer
Domain 1: Bias arising from the randomisation process	1. 1. Was the allocation sequence random?	Yes

Section	Question	Answer
Domain 1: Bias arising from the randomisation process	1. 2. Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Yes
Domain 1: Bias arising from the randomisation process	1.3 Did baseline differences between intervention groups suggest a problem with the randomisation process?	No
Domain 1: Bias arising from the randomisation process	Risk of bias judgement for the randomisation process	Low
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.1. Were participants aware of their assigned intervention during the trial?	No
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	No
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	Not applicable
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.4. If Y/PY to 2.3: Were these deviations from intended intervention balanced between groups?	Not applicable
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.5 If N/PN/NI to 2.4: Were these deviations likely to have affected the outcome?	Not applicable

Section	Question	Answer
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Yes
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	Not applicable
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	Risk of bias for deviations from the intended interventions (effect of assignment to intervention)	Low
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.1. Were participants aware of their assigned intervention during the trial?	No
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	No
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.3. If Y/PY/NI to 2.1 or 2.2: Were important co-interventions balanced across intervention groups?	Probably yes
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.4. Could failures in implementing the intervention have affected the outcome?	Probably no
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.5. Did study participants adhere to the assigned intervention regimen?	Probably yes

Section	Question	Answer
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.6. If N/PN/NI to 2.3 or 2.5 or Y/PY/NI to 2.4: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	Yes
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	Risk of bias judgement for deviations from the intended interventions (effect of adhering to intervention)	Low
Domain 3. Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomised?	Yes
Domain 3. Bias due to missing outcome data	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	Not applicable
Domain 3. Bias due to missing outcome data	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Not applicable
Domain 3. Bias due to missing outcome data	3.4 If Y/PY/NI to 3.3: Do the proportions of missing outcome data differ between intervention groups?	Not applicable
Domain 3. Bias due to missing outcome data	3.5 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	Not applicable
Domain 3. Bias due to missing outcome data	Risk-of-bias judgement for missing outcome data	Low
Domain 4. Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	No
Domain 4. Bias in measurement of the outcome	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups ?	No

Section	Question	Answer
Domain 4. Bias in measurement of the outcome	4.3 If N/PN/NI to 4.1 and 4.2: Were outcome assessors aware of the intervention received by study participants ?	No
Domain 4. Bias in measurement of the outcome	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	No
Domain 4. Bias in measurement of the outcome	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	Not applicable
Domain 4. Bias in measurement of the outcome	Risk-of-bias judgement for measurement of the outcome	Low
Domain 5. Bias in selection of the reported result	5.1 Was the trial analysed in accordance with a pre-specified plan that was finalised before unblinded outcome data were available for analysis ?	Yes
Domain 5. Bias in selection of the reported result	5.2 Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	No/Probably no
Domain 5. Bias in selection of the reported result	5.3 Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple analyses of the data?	No/Probably no
Domain 5. Bias in selection of the reported result	Risk-of-bias judgement for selection of the reported result	Low
Overall bias and Directness	Risk of bias judgement	Low
Overall bias and Directness	Overall Directness	Directly applicable

Pollock, 2019

Bibliographic Reference	Pollock, Carol; Stefansson, Bergur; Reyner, Daniel; Rossing, Peter; Sjostrom, C David; Wheeler, David C; Langkilde, Anna Maria; Heerspink, Hiddo J L; Albuminuria-lowering effect of dapagliflozin alone and in combination with saxagliptin and effect of dapagliflozin and saxagliptin on glycaemic control in patients with type 2 diabetes and chronic kidney disease (DELIGHT): a randomised, double-blind, placebo-controlled trial.; The lancet. Diabetes & endocrinology; 2019; vol. 7 (no. 6); 429-441

Study details

Secondary publication of another included study- see primary study for details	
Other publications associated with this study included in review	
Trial registration number and/or trial name	DELIGHT study - NCT02547935
Study type	Randomised controlled trial (RCT)
Study location	Australia, Canada, Japan, South Korea, Mexico, South Africa, Spain, Taiwan, and US
Study setting	116 research centres
Study dates	2015 - 2018

Sources of funding	AstraZeneca
Inclusion criteria	CKD eGFR 20 to 80 mL/min per 1.73 m² to enter the lead-in period (25 to 75 mL/min per 1.73 m² for randomisation) Albuminuria Urinary albumin/creatinine ratio 30 to 3500 mg/ Age 18 years or older Diabetes type 2 diabetes for more than 12 months Other HbA1c of 7.0 to 11.0% (53 to 97 mmol/mol) at screening Treatment stable glucose-lowering and antihypertensive treatments, including angiotensin-converting enzyme inhibitors or angiotensin II receptor blockers, at a clinically appropriate dose for at least 12 weeks before randomisation
Exclusion criteria	Other conditions type 1 diabetes, known non-diabetic kidney disease, severe cardiovascular disease, two or more major hypoglycaemia events within 12 weeks before screening, haemoglobin less than 9 g/dL (or 5.6 mmol/L), or evidence of hepatic disease, poorly controlled blood pressure (systolic blood pressure ≥180 mm Hg or diastolic blood pressure ≥110 mm Hg) Treatment current use of SGLT2 inhibitors, GLP-1 receptor agonists, or DPP-4 inhibitors, and long-term treatment with glucocorticoids

Intervention(s)	Antihypertensive treatments were to be kept stable throughout the entire study, from the start of the run-in period until the end of follow-up
Outcome measures	Adverse events Minor hypoglycaemia was defined as symptomatic episodes with capillary or plasma glucose <3.5 mmol/L [63 mg/dL], regardless of need for external assistance; or asymptomatic capillary or plasma glucose <3.5 mmol/L [63 mg/dL] not qualifying as a major episode. Major hypoglycaemia was defined as symptomatic episodes requiring external [third party] assistance because of severe impairment in consciousness or behaviour [capillary or plasma glucose <3 mmol/L or <54 mg/dL] and prompt recovery after glucose or glucagon administration. Safety endpoints assessed in this study were the change from baseline in eGFR at week 24 and at week 27 (3 weeks after treatment completion), and the proportion of patients who discontinued study medication because of a sustained increase in serum creatinine of at least 1·5 times from baseline concentration. Safety endpoints were assessed in the safety analysis set, which comprised all patients who received at least one dose of double-blind study drug during the 24-week double-blind treatment period. Albuminuria 24-h urinary albumin excretion (g/day) reported as median at baseline and 24 weeks and adjusted mean change from baseline Mortality 1) all cause Urine Albumin Creatinine Ratio 1) The primary efficacy endpoint for the dapagliflozin treatment group was percentage change in UACR at week 24 versus baseline. 2) proportion of patients achieving a reduction of more than 30% in UACR 3) post-hoc analysis of change in UACR for patients achieving or not achieving each individual component of the clinical benefit endpoint was also done HbA1c For the dapagliflozin–saxagliptin combined treatment group, percentage change in HbA1c at week 24 versus baseline and percentage change in UACR at week 24 versus baseline were the coprimary endpoints. Because the HbA1c-lowering effect

	of dapagliflozin is attenuated in patients with kidney impairment, change in HbA1c was not a coprimary endpoint but a secondary endpoint in the dapagliflozin-only treatment group. 2) proportion of patients achieving a reduction in HbA1c to less than 7·0% (53 mmol/mol) at week 24. 3) the proportion of patients achieving a clinical benefit endpoint (ie, responders), defined as reductions of at least 0·3% (3·3 mmol/mol) in HbA1c, at least 3% in bodyweight, and at least 3 mm Hg in seated systolic blood pressure.
Number of participants	461
Duration of follow-up	24 weeks
Loss to follow-up	Dapagliflozin plus Saxagliptin = 7 Dapagliflozin = 14 Placebo = 10
Methods of analysis	Intention to treat
Additional comments	

Study arms

Dapagliflozin plus Saxagliptin (N = 157)

Outcome measures	Adverse events Minor hypoglycaemia was defined as symptomatic episodes with capillary or plasma glucose <3.5 mmol/L [63 mg/dL], regardless of need for external assistance; or asymptomatic	Adverse events Minor hypoglycaemia was defined as symptomatic episodes with capillary or plasma glucose <3.5 mmol/L [63 mg/dL], regardless of need for external assistance; or asymptomatic

	capillary or plasma glucose <3.5 mmol/L [63 mg/dL] not qualifying as a major episode. Major hypoglycaemia was defined as symptomatic episodes requiring external [third party] assistance because of severe impairment in consciousness or behaviour [capillary or plasma glucose <3 mmol/L or <54 mg/dL] and prompt recovery after glucose or glucagon administration. Albuminuria 24-h urinary albumin excretion (g/day) reported as median at baseline and 24 weeks and adjusted mean change from baseline Mortality 1) all cause	capillary or plasma glucose <3.5 mmol/L [63 mg/dL] not qualifying as a major episode. Major hypoglycaemia was defined as symptomatic episodes requiring external [third party] assistance because of severe impairment in consciousness or behaviour [capillary or plasma glucose <3 mmol/L or <54 mg/dL] and prompt recovery after glucose or glucagon administration. Albuminuria 24-h urinary albumin excretion (g/day) reported as median at baseline and 24 weeks and adjusted mean change from baseline Mortality 1) all cause

once-daily dapagliflozin (10 mg) and saxagliptin (2·5 mg)

Dapagliflozin (N = 151)

Outcome measures	Adverse events Minor hypoglycaemia was defined as symptomatic episodes with capillary or plasma glucose <3.5 mmol/L [63 mg/dL], regardless of need for external assistance; or asymptomatic capillary or plasma glucose <3.5 mmol/L [63 mg/dL] not qualifying as a major episode. Major hypoglycaemia was defined as symptomatic episodes requiring external [third party] assistance because of severe impairment in consciousness or behaviour [capillary or plasma glucose <3 mmol/L or <54 mg/dL] and prompt recovery after glucose or glucagon administration.	Adverse events Minor hypoglycaemia was defined as symptomatic episodes with capillary or plasma glucose <3.5 mmol/L [63 mg/dL], regardless of need for external assistance; or asymptomatic capillary or plasma glucose <3.5 mmol/L [63 mg/dL] not qualifying as a major episode. Major hypoglycaemia was defined as symptomatic episodes requiring external [third party] assistance because of severe impairment in consciousness or behaviour [capillary or plasma glucose <3 mmol/L or <54 mg/dL] and prompt recovery after glucose or glucagon administration.

	Albuminuria 24-h urinary albumin excretion (g/day) reported as median at baseline and 24 weeks and adjusted mean change from baseline Mortality 1) all cause	Albuminuria 24-h urinary albumin excretion (g/day) reported as median at baseline and 24 weeks and adjusted mean change from baseline Mortality 1) all cause

once-daily dapagliflozin (10 mg)

Placebo (N = 153)

Outcome measures	Adverse events Minor hypoglycaemia was defined as symptomatic episodes with capillary or plasma glucose <3.5 mmol/L [63 mg/dL], regardless of need for external assistance; or asymptomatic capillary or plasma glucose <3.5 mmol/L [63 mg/dL] not qualifying as a major episode. Major hypoglycaemia was defined as symptomatic episodes requiring external [third party] assistance because of severe impairment in consciousness or behaviour [capillary or plasma glucose <3 mmol/L or <54 mg/dL] and prompt recovery after glucose or glucagon administration. Albuminuria 24-h urinary albumin excretion (g/day) reported as median at baseline and 24 weeks and adjusted mean change from baseline Mortality	Adverse events Minor hypoglycaemia was defined as symptomatic episodes with capillary or plasma glucose <3.5 mmol/L [63 mg/dL], regardless of need for external assistance; or asymptomatic capillary or plasma glucose <3.5 mmol/L [63 mg/dL] not qualifying as a major episode. Major hypoglycaemia was defined as symptomatic episodes requiring external [third party] assistance because of severe impairment in consciousness or behaviour [capillary or plasma glucose <3 mmol/L or <54 mg/dL] and prompt recovery after glucose or glucagon administration. Albuminuria 24-h urinary albumin excretion (g/day) reported as median at baseline and 24 weeks and adjusted mean change from baseline Mortality

	1) all cause	1) all cause

once-daily matched placebo

Characteristics

Arm-level characteristics

Characteristic	Dapagliflozin plus Saxagliptin (N = 157)	Dapagliflozin (N = 151)	Placebo (N = 153)
Age (years) Mean (SD)	64 (9.2)	64.7 (8.6)	64.7 (8.5)
Female Sample size	n = 45 ; % = 29	n = 43 ; % = 30	n = 43 ; % = 29
White Sample size	n = 64 ; % = 43	n = 55 ; % = 38	n = 64 ; % = 43
Black Sample size	n = 11 ; % = 7	n = 7 ; % = 5	n = 11 ; % = 7
Asian Sample size	n = 53 ; % = 36	n = 67 ; % = 46	n = 53 ; % = 36
Other Sample size	n = 20 ; % = 14	n = 16 ; % = 11	n = 20 ; % = 14

Characteristic	Dapagliflozin plus Saxagliptin (N = 157)	Dapagliflozin (N = 151)	Placebo (N = 153)
BMI (kg/m²) Mean (SD)	30.81 (5.4)	30.19 (5.3)	30.81 (5.4)
eGFR (mL/min per 1·73 m²) Mean (SD)	49 (13)	50.2 (13)	47.7 (13.5)
Urine Albumin to Creatinine Ratio (mg/g) Median (IQR)	218.4 (74 to 936)	270 (69 to 751)	257.5 (80 to 949)
Normoalbuminuria Sample size	n = 12 ; % = 8	n = 10 ; % = 7	n = 11 ; % = 7
Microalbuminuria Sample size	n = 73 ; % = 47	n = 64 ; % = 44	n = 65 ; % = 44
Macroalbuminuria Sample size	n = 70 ; % = 45	n = 71 ; % = 49	n = 72 ; % = 49
Serum Creatinine (mg/dL) Mean (SD)	1.4 (0.4)	1.4 (0.3)	1.4 (0.4)

Critical appraisal - GUT Cochrane Risk of Bias tool (RoB 2.0) Normal RCT

Section	Question	Answer
Domain 1: Bias arising from the randomisation process	1. 1. Was the allocation sequence random?	Yes
Domain 1: Bias arising from the randomisation process	1. 2. Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Yes
Domain 1: Bias arising from the randomisation process	1.3 Did baseline differences between intervention groups suggest a problem with the randomisation process?	No
Domain 1: Bias arising from the randomisation process	Risk of bias judgement for the randomisation process	Low
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.1. Were participants aware of their assigned intervention during the trial?	No
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	No
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	Not applicable
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.4. If Y/PY to 2.3: Were these deviations from intended intervention balanced between groups?	Not applicable
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.5 If N/PN/NI to 2.4: Were these deviations likely to have affected the outcome?	Not applicable

Section	Question	Answer
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Yes
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	Not applicable
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	Risk of bias for deviations from the intended interventions (effect of assignment to intervention)	Low
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.1. Were participants aware of their assigned intervention during the trial?	No
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	No
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.3. If Y/PY/NI to 2.1 or 2.2: Were important co-interventions balanced across intervention groups?	Not applicable
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.4. Could failures in implementing the intervention have affected the outcome?	Probably no
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.5. Did study participants adhere to the assigned intervention regimen?	Yes

Section	Question	Answer
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.6. If N/PN/NI to 2.3 or 2.5 or Y/PY/NI to 2.4: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	Not applicable
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	Risk of bias judgement for deviations from the intended interventions (effect of adhering to intervention)	Low
Domain 3. Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomised?	Yes
Domain 3. Bias due to missing outcome data	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	Not applicable
Domain 3. Bias due to missing outcome data	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Not applicable
Domain 3. Bias due to missing outcome data	3.4 If Y/PY/NI to 3.3: Do the proportions of missing outcome data differ between intervention groups?	Not applicable
Domain 3. Bias due to missing outcome data	3.5 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	Not applicable
Domain 3. Bias due to missing outcome data	Risk-of-bias judgement for missing outcome data	Low
Domain 4. Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	No
Domain 4. Bias in measurement of the outcome	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups ?	Probably no

Section	Question	Answer
Domain 4. Bias in measurement of the outcome	4.3 If N/PN/NI to 4.1 and 4.2: Were outcome assessors aware of the intervention received by study participants ?	No
Domain 4. Bias in measurement of the outcome	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Not applicable
Domain 4. Bias in measurement of the outcome	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	Not applicable
Domain 4. Bias in measurement of the outcome	Risk-of-bias judgement for measurement of the outcome	Low
Domain 5. Bias in selection of the reported result	5.1 Was the trial analysed in accordance with a pre-specified plan that was finalised before unblinded outcome data were available for analysis ?	Yes
Domain 5. Bias in selection of the reported result	5.2 Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	No/Probably no
Domain 5. Bias in selection of the reported result	5.3 Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple analyses of the data?	No/Probably no
Domain 5. Bias in selection of the reported result	Risk-of-bias judgement for selection of the reported result	Low
Overall bias and Directness	Risk of bias judgement	Low
Overall bias and Directness	Overall Directness	Directly applicable

Wanner, 2018

Bibliographic Reference	Wanner, Christoph; Lachin, John M; Inzucchi, Silvio E; Fitchett, David; Mattheus, Michaela; George, Jyothis; Woerle, Hans J; Broedl, Uli C; von Eynatten, Maximilian; Zinman, Bernard; EMPA-REG OUTCOME, Investigators; Empagliflozin and Clinical Outcomes in Patients With Type 2 Diabetes Mellitus, Established Cardiovascular Disease, and Chronic Kidney Disease.; Circulation; 2018; vol. 137 (no. 2); 119-129

Study details

Secondary publication of another included study- see primary study for details	
Other publications associated with this study included in review	
Trial registration number and/or trial name	EMPA - REG: ClinicalTrials.gov number, NCT01131676
Study type	Randomised controlled trial (RCT)
Study location	42 countries - not specified
Study setting	590 sites - North America [plus Australia and New Zealand], Latin America, Europe, Africa, or Asia
Study dates	Randomization from September 2010 through April 2013; date for last data collection point and follow-up not outlined

Sources of funding	Supported by Boehringer Ingelheim and Eli Lilly
Inclusion criteria	CKD estimated glomerular filtration rate (eGFR) of at least 30 ml per minute per 1.73 m2 of body-surface area, according to the Modification of Diet in Renal Disease criteria. With prevalent kidney disease (defined as eGFR <60 mL·min−1·1.73 m−2 and/or urine albumin-creatinine ratio >300 mg/g) at baseline Age Adults (aged 18 years and older) Diabetes type 2 diabetes
Exclusion criteria	Other conditions Cancer Liver disease Treatment Received glucose-lowering agents for at least 12 weeks before randomization and had a glycated haemoglobin level of at <7.0% and > than 10.0%. No glucose-lowering agents for at least 12 weeks before randomization. HbA1c glycated haemoglobin level of at <7.0% and > 9.0% Pregnant or breastfeeding not using adequate contraception

Intervention(s)	Empagliflozin 10 mg (n=2345 - total population) or 25 mg (n=2342 - total population) of empagliflozin
Comparator	
Outcome measures	Adverse events Safety was assessed by evaluation of adverse events in subgroups stratified by eGFR <45, 45 to <60, and ≥60 mL·min−1·1.73 m−2 at baseline and are depicted for empagliflozin pooled, empagliflozin 10 mg, empagliflozin 25 mg, and placebo. Confirmed hypoglycemia adverse events were documented episodes with plasma glucose ≤70 mg/dL and/or requiring assistance. Events consistent with urinary tract infection, genital infection, acute renal failure, volume depletion, bone fracture, and hyperkalemia were based on searches of adverse events reported by investigators. From the trial database, authors identified events consistent with lower limb amputation from events reported as adverse events, from those reported as a “medical procedure” in electronic case report forms or in investigator comments describing adverse events, and via a systematic search of serious adverse event narratives. Composite cardiovascular outcome 3-point MACE: composite of nonfatal stroke, nonfatal myocardial infarction, and cardiovascular death Mortality Cardiovascular-related mortality All-cause mortality Heart failure Hospitalization for heart failure Myocardial Infarction and hospitalization for unstable angina Stroke

	Stroke, or atherosclerotic disease; Fatal or nonfatal stroke; Nonfatal stroke Hospitalisation All-cause hospitalization
Number of participants	7028 patients underwent randomization; 7020 were treated and included in the primary analysis. 2250 were included in the sub analysis of those with renal failure.
Duration of follow-up	3.1 years (mean)
Loss to follow-up	In the total population - 8/7020 randomized were not included in the primary analysis (0.1%). 97.0% of patients completed the study (n=6809), with 25.4% of patients prematurely discontinuing a study drug (n=1780). Final vital status was available for 99.2% of patients (n=6967).
Methods of analysis	modified Intention to treat Cox proportional-hazards model, with study group, age, sex, baseline body-mass index, baseline glycated haemoglobin level, baseline eGFR, and geographic region as factors; Kaplan–Meier estimates for death from any cause.
Additional comments	

Study arms

Empagliflozin (N = 1498)

Subgroup with eGFR (Modification of Diet in Renal Disease) <60 mL·min−1·1.73 m−2 and/or macroalbuminuria (urine albumin-creatinine ratio >300 mg/g)

Placebo (N = 752)

Subgroup with eGFR (Modification of Diet in Renal Disease) <60 mL·min−1·1.73 m−2 and/or macroalbuminuria (urine albumin-creatinine ratio >300 mg/g)

Characteristics

Arm-level characteristics

Characteristic	**Empagliflozin (N = 1498)**	**Placebo (N = 752)**
Age (years) Mean (SD)	66.2 (8)	66 (8.5)
Male Sample size	n = 1033 ; % = 69	n = 529 ; % = 70.3
White Sample size	n = 1070 ; % = 71.4	n = 544 ; % = 72.3
Asian Sample size	n = 338 ; % = 22.6	n = 167 ; % = 22.2
Black/African-American Sample size	n = 74 ; % = 4.9	n = 34 ; % = 4.5
Other/missing Sample size	n = 16 ; % = 1.1	n = 7 ; % = 0.9

Characteristic	Empagliflozin (N = 1498)	Placebo (N = 752)
Not hispanic or latino Sample size	n = 1204 ; % = 80.4	n = 617 ; % = 82
Hispanic or Latino Sample size	n = 293 ; % = 19.6	n = 134 ; % = 17.8
MIssing Sample size	n = 1 ; % = 0.1	n = 1 ; % = 0.1
BMI Mean (SD)	30.8 (5.4)	30.8 (5.4)
Estimated glomerular filtration rate (MDRD), mL/min/1.73m2 Mean (SD)	54.5 (16.1)	54.3 (15.2)
<30 mg/g Sample size	n = 566 ; % = 37.8	n = 283 ; % = 37.6
30 to 300 mg/g Sample size	n = 411 ; % = 27.4	n = 205 ; % = 27.3
>300 mg/g Sample size	n = 509 ; % = 34	n = 260 ; % = 34.6
MIssing	n = 12 ; % = 0.8	n = 4 ; % = 0.5

Characteristic	Empagliflozin (N = 1498)	Placebo (N = 752)
Sample size		

Critical appraisal - GUT Cochrane Risk of Bias tool (RoB 2.0) Normal RCT

Section	Question	Answer
Domain 1: Bias arising from the randomisation process	1. 1. Was the allocation sequence random?	Yes
Domain 1: Bias arising from the randomisation process	1. 2. Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Probably yes
Domain 1: Bias arising from the randomisation process	1.3 Did baseline differences between intervention groups suggest a problem with the randomisation process?	Yes
Domain 1: Bias arising from the randomisation process	Risk of bias judgement for the randomisation process	Low *((Randomized, double-blind, placebo-controlled trial. Randomization process outlined, but protocol for allocation concealment not specified. No significant differences outlined for baseline characteristics post randomisation))*
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.1. Were participants aware of their assigned intervention during the trial?	No

Section	Question	Answer
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	No
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	Not applicable
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.4. If Y/PY to 2.3: Were these deviations from intended intervention balanced between groups?	Not applicable
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.5 If N/PN/NI to 2.4: Were these deviations likely to have affected the outcome?	Not applicable
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Yes
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	Not applicable

Section	Question	Answer
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	Risk of bias for deviations from the intended interventions (effect of assignment to intervention)	Low *((Randomized, double-blind, placebo-controlled trial. Randomization process outlined, but protocol for allocation concealment not specified. mITT undertaken for primary analysis))*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.1. Were participants aware of their assigned intervention during the trial?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.3. If Y/PY/NI to 2.1 or 2.2: Were important co-interventions balanced across intervention groups?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.4. Could failures in implementing the intervention have affected the outcome?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.5. Did study participants adhere to the assigned intervention regimen?	*N/A.*

Section	Question	Answer
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.6. If N/PN/NI to 2.3 or 2.5 or Y/PY/NI to 2.4: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	Risk of bias judgement for deviations from the intended interventions (effect of adhering to intervention)	*N/A.*
Domain 3. Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomised?	Yes
Domain 3. Bias due to missing outcome data	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	Not applicable
Domain 3. Bias due to missing outcome data	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Not applicable
Domain 3. Bias due to missing outcome data	3.4 If Y/PY/NI to 3.3: Do the proportions of missing outcome data differ between intervention groups?	Not applicable
Domain 3. Bias due to missing outcome data	3.5 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	Not applicable
Domain 3. Bias due to missing outcome data	Risk-of-bias judgement for missing outcome data	Low *((Data presented for n=7020 participants for all outcomes except silent myocardial infarction (n=3589); 97.0% of patients completed*

Section	Question	Answer
		the study (n=6809), with 25.4% of patients prematurely discontinuing a study drug (n=1780). Final vital status was available for 99.2% of patients (n=6967).))
Domain 4. Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	No
Domain 4. Bias in measurement of the outcome	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups ?	No
Domain 4. Bias in measurement of the outcome	4.3 If N/PN/NI to 4.1 and 4.2: Were outcome assessors aware of the intervention received by study participants ?	No
Domain 4. Bias in measurement of the outcome	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Not applicable
Domain 4. Bias in measurement of the outcome	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	Not applicable
Domain 4. Bias in measurement of the outcome	Risk-of-bias judgement for measurement of the outcome	Low *((Study outlined as a double-blind randomized controlled trial. Clinical event rates were the measures for the primary outcome and secondary outcome. Definitions of major clinical outcomes prespecified. Cardiovascular outcome events and deaths were prospectively adjudicated by two clinical-events committees.))*

Section	Question	Answer
Domain 5. Bias in selection of the reported result	5.1 Was the trial analysed in accordance with a pre-specified plan that was finalised before unblinded outcome data were available for analysis ?	Yes
Domain 5. Bias in selection of the reported result	5.2 Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	No/Probably no
Domain 5. Bias in selection of the reported result	5.3 Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple analyses of the data?	No/Probably no
Domain 5. Bias in selection of the reported result	Risk-of-bias judgement for selection of the reported result	Low *((Evidence of prespecified analytical plan; Outcomes reported for most participants against prespecified and clearly defined outcomes using clinical event rates.))*
Overall bias and Directness	Risk of bias judgement	Low
Overall bias and Directness	Overall Directness	Directly applicable

Wheeler, 2021

Bibliographic Reference	Wheeler, David C; Stefansson, Bergur V; Jongs, Niels; Chertow, Glenn M; Greene, Tom; Hou, Fan Fan; McMurray, John J V; Correa-Rotter, Ricardo; Rossing, Peter; Toto, Robert D; Sjostrom, C David; Langkilde, Anna Maria; Heerspink, Hiddo J L; DAPA-CKD Trial Committees and, Investigators; Effects of dapagliflozin on major adverse kidney and cardiovascular events in

patients with diabetic and non-diabetic chronic kidney disease: a prespecified analysis from the DAPA-CKD trial.; The lancet. Diabetes & endocrinology; 2021; vol. 9 (no. 1); 22-31

Study details

Trial registration number and/or trial name	DAPA-CKD: NCT03036150
Study type	Randomised controlled trial (RCT)
Study location	21 countries (Argentina, Brazil, Canada, China, Denmark, Germany, Hungary, India, Japan, Mexico, Peru, Philippines, Poland, Russia, South Korea, Spain, Sweden, UK, Ukraine, USA, and Vietnam).
Study setting	386 study sites
Study dates	between Feb 2, 2017, and June 12, 2020.
Sources of funding	AstraZeneca
Inclusion criteria	CKD chronic kidney disease, defined as an eGFR between 25 and 75 mL/min/1·73 m² and a urinary albumin-to-creatinine ratio (UACR) between 200 and 5000 mg/g (22·6 to 565·6 mg/mmol). Treatment All participants were required to be receiving a stable dose of an ACE inhibitor or ARB for at least 4 weeks before enrolment into the trial, unless contraindicated.
Exclusion criteria	Other conditions lupus nephritis, or anti-neutrophil cytoplasmic antibody associated vasculitis

	Treatment Participants receiving immunotherapy for primary or secondary kidney disease within the 6 months before enrolment were also excluded Diabetes diagnosis of type 1 diabetes, Renal polycystic kidney disease
Intervention(s)	Dapagliflozin
Outcome measures	Composite kidney outcome The primary outcome of the trial was a composite of a sustained decline of 50% or more in eGFR (confirmed by a second serum creatinine after at least 28 days), onset of end-stage kidney disease (defined as maintenance dialysis for more than 28 days, kidney transplantation, or eGFR <15 mL/min per 1·73 m² confirmed by a second measurement after at least 28 days), or death from kidney or cardiovascular causes. 2) a kidney-specific composite outcome defined in the same way as the primary outcome but excluding cardiovascular death 3) chronic dialysis, kidney transplantation, and death from kidney-related causes. Adverse events These data included all serious adverse events, all adverse events leading to discontinuation, and specified adverse events of interest (amputations, potential diabetic ketoacidosis, bone fractures, kidney-related adverse events, major hypoglycaemia, and symptoms of volume depletion). Events of potential diabetic ketoacidosis were adjudicated by an

	independent adjudication committee. Serious adverse events and discontinuations related to urinary tract infections and genital infections were also reported Composite cardiovascular outcome composite outcome of cardiovascular death or hospital admission for heart failure Mortality all-cause mortality
Number of participants	4304 (2152 to dapagliflozin and 2152 to placebo). 2906 (68%) participants had type 2 diabetes and 1398 (32%) did not.
Duration of follow-up	followed up for a median of 2·4 years (IQR 2·0–2·7).
Loss to follow-up	Not reported (all data analyzed via intention to treat)
Methods of analysis	Intention to treat Authors fitted a Cox proportional-hazards regression model, stratified by type 2 diabetes and UACR and adjusted for baseline eGFR, to estimate the hazard ratio (HR) and 95% CIs for dapagliflozin compared with placebo in participants with or without type 2 diabetes, and within each prespecified subgroup based on reported cause of chronic kidney disease. Logistic regression was used to estimate the odds ratio and 95% CI for dapagliflozin compared with placebo in participants with and without type 2 diabetes for safety data.

Study arms

dapagliflozin 10 mg (N = 1455)

dapagliflozin 10 mg (AstraZeneca, Gothenburg, Sweden) once daily

Placebo (N = 1451)

matching placebo

Characteristics

Arm-level characteristics

Characteristic	dapagliflozin 10 mg (N = 1455)	Placebo (N = 1451)
Age Mean (SD)	64.1 (9.8)	64.7 (9.5)
Female Sample size	n = 494 ; % = 34	n = 471 ; % = 32
White Sample size	n = 751 ; % = 52	n = 790 ; % = 54
Black or African American Sample size	n = 76 ; % = 5	n = 61 ; % = 4
Asian Sample size	n = 481 ; % = 33	n = 451 ; % = 31

Characteristic	dapagliflozin 10 mg (N = 1455)	Placebo (N = 1451)
Other Sample size	n = 147 ; % = 10	n = 149 ; % = 10
eGFR ml/min 1.73m^2 Mean (SD)	44 (12.6)	43.6 (12.6)
Median UACR mg/g Median (IQR)	1024.5 (472.5 to 2111)	1004.5 (493.3 to 2017)

Critical appraisal - GUT Cochrane Risk of Bias tool (RoB 2.0) Normal RCT

Section	Question	Answer
Domain 1: Bias arising from the randomisation process	1. 1. Was the allocation sequence random?	Yes
Domain 1: Bias arising from the randomisation process	1. 2. Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Yes
Domain 1: Bias arising from the randomisation process	1.3 Did baseline differences between intervention groups suggest a problem with the randomisation process?	No
Domain 1: Bias arising from the randomisation process	Risk of bias judgement for the randomisation process	Low

Section	Question	Answer
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.1. Were participants aware of their assigned intervention during the trial?	No
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	No
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	No/Probably no
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.4. If Y/PY to 2.3: Were these deviations from intended intervention balanced between groups?	Not applicable
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.5 If N/PN/NI to 2.4: Were these deviations likely to have affected the outcome?	Not applicable
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Yes

Section	Question	Answer
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	Not applicable
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	Risk of bias for deviations from the intended interventions (effect of assignment to intervention)	Low *(deviations from intended treatment appeared to be low (see supplement), intent to treat analysis used)*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.1. Were participants aware of their assigned intervention during the trial?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.3. If Y/PY/NI to 2.1 or 2.2: Were important co-interventions balanced across intervention groups?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.4. Could failures in implementing the intervention have affected the outcome?	*N/A.*

Section	Question	Answer
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.5. Did study participants adhere to the assigned intervention regimen?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.6. If N/PN/NI to 2.3 or 2.5 or Y/PY/NI to 2.4: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	Risk of bias judgement for deviations from the intended interventions (effect of adhering to intervention)	*N/A.*
Domain 3. Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomised?	Yes
Domain 3. Bias due to missing outcome data	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	Not applicable
Domain 3. Bias due to missing outcome data	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Not applicable
Domain 3. Bias due to missing outcome data	3.4 If Y/PY/NI to 3.3: Do the proportions of missing outcome data differ between intervention groups?	Not applicable
Domain 3. Bias due to missing outcome data	3.5 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	Not applicable

Section	Question	Answer
Domain 3. Bias due to missing outcome data	Risk-of-bias judgement for missing outcome data	Low *(data available for over 99% of participants (as reported in trial supplement))*
Domain 4. Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	No
Domain 4. Bias in measurement of the outcome	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups ?	Probably no
Domain 4. Bias in measurement of the outcome	4.3 If N/PN/NI to 4.1 and 4.2: Were outcome assessors aware of the intervention received by study participants ?	No
Domain 4. Bias in measurement of the outcome	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Not applicable
Domain 4. Bias in measurement of the outcome	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	Not applicable
Domain 4. Bias in measurement of the outcome	Risk-of-bias judgement for measurement of the outcome	Low
Domain 5. Bias in selection of the reported result	5.1 Was the trial analysed in accordance with a pre-specified plan that was finalised before unblinded outcome data were available for analysis ?	Yes
Domain 5. Bias in selection of the reported result	5.2 Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	No/Probably no

Section	Question	Answer
Domain 5. Bias in selection of the reported result	5.3 Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple analyses of the data?	No/Probably no
Domain 5. Bias in selection of the reported result	Risk-of-bias judgement for selection of the reported result	Low *("An independent event adjudication committee adjudicated all clinical outcome events using rigorous prespecified endpoint definitions.")*
Overall bias and Directness	Risk of bias judgement	Low
Overall bias and Directness	Overall Directness	Directly applicable

Wiviott, 2019

Bibliographic Reference	Wiviott, Stephen D; Raz, Itamar; Bonaca, Marc P; Mosenzon, Ofri; Kato, Eri T; Cahn, Avivit; Silverman, Michael G; Zelniker, Thomas A; Kuder, Julia F; Murphy, Sabina A; Bhatt, Deepak L; Leiter, Lawrence A; McGuire, Darren K; Wilding, John P H; Ruff, Christian T; Gause-Nilsson, Ingrid A M; Fredriksson, Martin; Johansson, Peter A; Langkilde, Anna-Maria; Sabatine, Marc S; DECLARE-TIMI 58, Investigators; Dapagliflozin and Cardiovascular Outcomes in Type 2 Diabetes.; The New England journal of medicine; 2019; vol. 380 (no. 4); 347-357

Study details

Other publications associated with this study included in review	Mosenzon O, Wiviott SD, Cahn A, Rozenberg A, Yanuv I, Goodrich EL, Murphy SA, Heerspink HJ, Zelniker TA, Dwyer JP, Bhatt DL. Effects of dapagliflozin on development and progression of kidney disease in patients with type 2 diabetes: an analysis from the DECLARE–TIMI 58 randomised trial. The lancet Diabetes & endocrinology. 2019 Aug 1;7(8):606-17.

Trial registration number and/or trial name	DECLARE–TIMI - NCT01730534
Study type	Randomised controlled trial (RCT)
Study location	33 countries
Study setting	882 sites
Study dates	April 25, 2013 to September 11, 2018
Sources of funding	AstraZeneca
Inclusion criteria	Age 40 years of age or older Diabetes type 2 diabetes Cardiovascular Established atherosclerotic cardiovascular disease (defined as clinically evident ischemic heart disease, ischemic cerebrovascular disease, or peripheral artery disease) or multiple risk factors for atherosclerotic cardiovascular disease (age ≥55 years for men or ≥60 years for women plus at least one of dyslipidaemia, hypertension, or current tobacco use) eGFR Subgroup analysis of eGFR: ≥ 90 ml/min/1.73 m2; 60 to <90 ml/min/1.73 m2; and <60 ml/min/1.73 m2 HbA1c glycated hemoglobin level of at least 6.5% but less than 12.0%,

	Creatine clearance a creatinine clearance of 60 ml or more per minute
Exclusion criteria	Other conditions Chronic cystitis and/or recurrent urinary tract infection Pregnant or breastfeeding Diabetes Diagnosis of Type 1 diabetes mellitus History of bladder cancer or history of radiation therapy to the lower abdomen or pelvis at any time
Intervention(s)	Dapagliflozin - The use of other glucose-lowering agents (other than an open-label SGLT2 inhibitor, pioglitazone, or rosiglitazone) was at the discretion of the treating physician.
Comparator	Placebo
Outcome measures	Composite kidney outcome 1) renal composite outcome, defined as a sustained decrease of 40% or more in estimated glomerular filtration rate (eGFR) — calculated by means of the Chronic Kidney Disease Epidemiology Collaboration equation - to less than 60 ml per minute per 1.73 m2 of body-surface area, new end-stage renal disease, or death from renal or cardiovascular causes. 2) A prespecified additional renal composite outcome included all the criteria described for the secondary renal outcome except for cardiovascular death eGFR Adverse events

	Serious adverse events and adverse events leading to discontinuation of dapagliflozin or placebo were collected Composite cardiovascular outcome 1) MACE (defined as cardiovascular death, myocardial infarction, or ischemic stroke). 2) a composite of cardiovascular death or hospitalization for heart failure Mortality death from any cause
Number of participants	Of the 17 160 participants who were randomly assigned, 8162 (47·6%) had an eGFR of at least 90 mL/min per 1·73 m², 7732 (45·1%) had an eGFR of 60 to less than 90 mL/min per 1·73 m², and 1265 (7·4%) had an eGFR of less than 60 mL/min per 1·73 m² at baseline (one participant had missing data for eGFR).
Duration of follow-up	Patients were followed for a median of 4.2 years (interquartile range, 3.9 to 4.4), for a total of 69,547 patient-years of follow-up.
Loss to follow-up	A total of 17,160 participants completed the run-in phase and were eligible to undergo randomization. A total of 3962 patients discontinued the trial regimen prematurely, at a rate of 5.7% per year, including 1811 of 8574 patients (21.1%) in the dapagliflozin group and 2151 of 8569 (25.1%) in the placebo group. Rates of withdrawal of consent (224 patients, at a rate of 0.3% per year) and loss to follow-up (30 patients, at a rate of <0.1% per year) were low and did not differ between the two groups. The primary analyses of cardiovascular safety and efficacy were performed with data from 17,160 patients who underwent randomization, with the exclusion of 30 participants from one site; data from patients at that site were excluded because of serious Good Clinical Practice violations in another trial that created uncertainty about the integrity of the data.
Methods of analysis	Intention-to-treat analysis Hazard ratios, 95% confidence intervals, and P values for time-to-event analyses are reported for the primary outcomes and were derived from a Cox proportional-hazards model in the overall population

Additional comments	

Study arms

dapagliflozin (N = 8582)

Outcome measures	Composite kidney outcome 1) renal composite outcome, defined as a sustained decrease of 40% or more in estimated glomerular filtration rate (eGFR) — calculated by means of the Chronic Kidney Disease Epidemiology Collaboration equation - to less than 60 ml per minute per 1.73 m2 of body-surface area, new end-stage renal disease, or death from renal or cardiovascular causes. 2) A prespecified additional renal composite outcome included all the criteria described for the secondary renal outcome except for cardiovascular death Adverse events Serious adverse events and adverse events leading to discontinuation of dapagliflozin or placebo were collected Composite cardiovascular outcome 1) MACE (defined as cardiovascular death, myocardial infarction, or ischemic stroke). 2) a composite of cardiovascular death or hospitalization for heart failure Mortality death from any cause
Loss to follow-up	A total of 17,160 participants completed the run-in phase and were eligible to undergo randomization. A total of 3962 patients discontinued the trial regimen prematurely, at a rate of 5.7% per year, including 1811 of 8574 patients (21.1%) in the dapagliflozin group and 2151 of 8569 (25.1%) in the placebo group. Rates of withdrawal of consent (224 patients, at a rate of 0.3% per year) and loss to follow-up (30 patients, at a rate of <0.1% per year) were low and did not differ between the two groups.

Methods of analysis	Intention-to-treat analysis

10 mg of dapagliflozin daily

Placebo (N = 8578)

Outcome measures	Composite kidney outcome 1) renal composite outcome, defined as a sustained decrease of 40% or more in estimated glomerular filtration rate (eGFR) — calculated by means of the Chronic Kidney Disease Epidemiology Collaboration equation - to less than 60 ml per minute per 1.73 m2 of body-surface area, new end-stage renal disease, or death from renal or cardiovascular causes. 2) A prespecified additional renal composite outcome included all the criteria described for the secondary renal outcome except for cardiovascular death Adverse events Serious adverse events and adverse events leading to discontinuation of dapagliflozin or placebo were collected Composite cardiovascular outcome 1) MACE (defined as cardiovascular death, myocardial infarction, or ischemic stroke). 2) a composite of cardiovascular death or hospitalization for heart failure Mortality death from any cause
Loss to follow-up	A total of 17,160 participants completed the run-in phase and were eligible to undergo randomization. A total of 3962 patients discontinued the trial regimen prematurely, at a rate of 5.7% per year, including 1811 of 8574 patients (21.1%) in the dapagliflozin group and 2151 of 8569 (25.1%) in the placebo group. Rates of withdrawal of consent (224 patients, at a rate of 0.3% per year) and loss to follow-up (30 patients, at a rate of <0.1% per year) were low and did not differ between the two groups.

matching placebo

Characteristics

Study-level characteristics

Characteristic	Study (N =)
eGFR ml/min/1.73 m2* **60 to <90** Sample size	n = 2866 ; % = 37.1
eGFR ml/min/1.73 m2 **<60** Sample size	n = 451 ; % = 35.7
eGFR ml/min/1.73 m2* **60 to <90** Mean (SD)	66.2 (6.5)
eGFR ml/min/1.73 m2 **<60** Mean (SD)	67.3 (6.6)
eGFR ml/min/1.73 m2* **60 to <90** Mean (SD)	32.1 (5.9)
eGFR ml/min/1.73 m2 **<60**	34.5 (6)

Characteristic	Study (N =)
Mean (SD)	
White eGFR ml/min/1.73 m2 **60 to <90** Sample size	n = 6313 ; % = 81.6
Non-White eGFR ml/min/1.73 m2 **60 to <90** Sample size	n = 1419 ; % = 18.4
White eGFR ml/min/1.73 m2 **<60** Sample size	n = 1088 ; % = 86
Non-white eGFR ml/min/1.73 m2 **<60** Sample size	n = 177 ; % = 14
eGFR ml/min/1.73 m2* **60 to <90** Mean (SD)	77 (8.5)
eGFR ml/min/1.73 m2 **<60** Mean (SD)	51.4 (7.2)
eGFR ml/min/1.73 m2* **60 to <90** Mean (SD)	8.1 (1.1)
eGFR ml/min/1.73 m2 **<60** Mean (SD)	8.2 (1.2)

Characteristic	Study (N =)
UACR < 30mg/g & eGFR ml/min/1.73 m2*** 60 to <90** Sample size	n = 5267 ; % = 69.5
UACR 30 – 300 mg/g & eGFR 60 to <90 Sample size	n = 1761 ; % = 23.2
UACR >300mg/g & eGFR 60 to <90 Sample size	n = 554 ; % = 7.3
UACR <30mg/g & eGFR <60 Sample size	n = 686 ; % = 55.6
UACR 30 – 300mg/g & eGFR <60 Sample size	n = 381 ; % = 30.9
UACR >300mg/g & eGFR <60 Sample size	n = 167 ; % = 13.5

Critical appraisal - GUT Cochrane Risk of Bias tool (RoB 2.0) Normal RCT

Section	Question	Answer
Domain 1: Bias arising from the randomisation process	1. 1. Was the allocation sequence random?	Yes
Domain 1: Bias arising from the randomisation process	1. 2. Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Yes
Domain 1: Bias arising from the randomisation process	1.3 Did baseline differences between intervention groups suggest a problem with the randomisation process?	No
Domain 1: Bias arising from the randomisation process	Risk of bias judgement for the randomisation process	Low
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.1. Were participants aware of their assigned intervention during the trial?	No
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	No
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	No/Probably no

Section	Question	Answer
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.4. If Y/PY to 2.3: Were these deviations from intended intervention balanced between groups?	Not applicable
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.5 If N/PN/NI to 2.4: Were these deviations likely to have affected the outcome?	Not applicable
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Yes
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	Not applicable
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	Risk of bias for deviations from the intended interventions (effect of assignment to intervention)	Low *(Intention to treat analysis used. In the total study population, a total of 3962 patients discontinued the trial regimen prematurely, at a rate of 5.7% per year, including 1811 of 8574 patients (21.1%) in the dapagliflozin group and 2151 of 8569 (25.1%) in the placebo group. This was unlikely to be related to study context.)*
Domain 2b: Risk of bias due to deviations from the intended	2.1. Were participants aware of their assigned intervention during the trial?	*N/A.*

Section	Question	Answer
interventions (effect of adhering to intervention)		
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.3. If Y/PY/NI to 2.1 or 2.2: Were important co-interventions balanced across intervention groups?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.4. Could failures in implementing the intervention have affected the outcome?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.5. Did study participants adhere to the assigned intervention regimen?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.6. If N/PN/NI to 2.3 or 2.5 or Y/PY/NI to 2.4: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	Risk of bias judgement for deviations from the intended interventions (effect of adhering to intervention)	*N/A.*

Section	Question	Answer
Domain 3. Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomised?	Yes
Domain 3. Bias due to missing outcome data	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	Not applicable
Domain 3. Bias due to missing outcome data	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Not applicable
Domain 3. Bias due to missing outcome data	3.4 If Y/PY/NI to 3.3: Do the proportions of missing outcome data differ between intervention groups?	Not applicable
Domain 3. Bias due to missing outcome data	3.5 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	Not applicable
Domain 3. Bias due to missing outcome data	Risk-of-bias judgement for missing outcome data	Low *(Rates of withdrawal of consent (224 patients, at a rate of 0.3% per year) and loss to follow-up (30 patients, at a rate of <0.1% per year) were low and did not differ between the two groups. (in the total study population))*
Domain 4. Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	Probably no
Domain 4. Bias in measurement of the outcome	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups ?	Probably no

Section	Question	Answer
Domain 4. Bias in measurement of the outcome	4.3 If N/PN/NI to 4.1 and 4.2: Were outcome assessors aware of the intervention received by study participants ?	No
Domain 4. Bias in measurement of the outcome	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Not applicable
Domain 4. Bias in measurement of the outcome	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	Not applicable
Domain 4. Bias in measurement of the outcome	Risk-of-bias judgement for measurement of the outcome	Low *(double-blind trial with predefined outcomes)*
Domain 5. Bias in selection of the reported result	5.1 Was the trial analysed in accordance with a pre-specified plan that was finalised before unblinded outcome data were available for analysis ?	Yes
Domain 5. Bias in selection of the reported result	5.2 Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	No/Probably no
Domain 5. Bias in selection of the reported result	5.3 Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple analyses of the data?	No/Probably no

Section	Question	Answer
Domain 5. Bias in selection of the reported result	Risk-of-bias judgement for selection of the reported result	Low
Overall bias and Directness	Risk of bias judgement	Low
Overall bias and Directness	Overall Directness	Directly applicable

Yale, 2013/ 2014

Bibliographic Reference	Yale, J-F; Bakris, G; Cariou, B; Nieto, J; David-Neto, E; Yue, D; Wajs, E; Figueroa, K; Jiang, J; Law, G; Usiskin, K; Meininger, G; DIA3004 Study, Group; Efficacy and safety of canagliflozin over 52 weeks in patients with type 2 diabetes mellitus and chronic kidney disease.; Diabetes, obesity & metabolism; 2014; vol. 16 (no. 10); 1016-27

Study details

Other publications associated with this study included in review	Yale JF, Bakris G, Cariou B, Yue D, David-Neto E, Xi L, Figueroa K, Wajs E, Usiskin K, Meininger G. Efficacy and safety of canagliflozin in subjects with type 2 diabetes and chronic kidney disease. Diabetes, Obesity and Metabolism. 2013 May;15(5):463-73.
Trial registration number and/or trial name	NCT01064414
Study type	Randomised controlled trial (RCT)
Study location	19 countries

Study setting	89 centres
Study dates	June 2010 to August 2012
Sources of funding	Janssen Research & Development, LLC
Inclusion criteria	Age aged ≥25 years Diabetes T2DM who had inadequate glycaemic control (HbA1c ≥7.0 and ≤10.5%) Treatment either not on AHA therapy or were on a stable AHA regimen (monotherapy or combination therapy with any approved agent including metformin, sulphonylurea, dipeptidyl peptidase-4 (DPP-4) inhibitor, α-glucosidase inhibitor, GLP-1 analogue, pioglitazone or insulin) for ≥8 weeks (≥12 weeks with pioglitazone) prior to the week –2 visit. Subjects on AHA regimens not consistent with local prescribing guidelines (e.g. metformin therapy) underwent an AHA adjustment period of up to 12 weeks before the placebo run-in period. Subjects were to remain on their stable AHA regimens through the completion of the 52-week treatment period (unless glycaemic rescue criteria were met). CKD stage 3 CKD (eGFR ≥30 and <50 ml/min/1.73 m2) eGFR Subjects were required to have generally stable renal function, as determined by a ≤25% decrease in eGFR from the screening to the week –2 visits.
Exclusion criteria	Diabetes

	Subjects were excluded if they had repeated fasting plasma glucose (FPG) >15.0 mmol/l (270 mg/dl) during the pretreatment phase; or a history of T1DM; Renal renal disease that required immunosuppressive therapy, dialysis or transplant; nephrotic syndrome or inflammatory renal disease; Cardiovascular New York Heart Association Class III-IV cardiovascular disease; myocardial infarction, unstable angina, revascularization procedure or cerebrovascular accident within 3 months prior to screening; or haemoglobin concentration <100 g/l (10 g/dl) at screening.
Intervention(s)	Dapagliflozin. During the double-blind, core treatment period, glycaemic rescue therapy (up-titration of current AHAs or step-wise addition of oral or non-oral AHAs) was initiated if FPG >15.0 mmol/l (270 mg/dl) after day 1 to week 6, >13.3 mmol/l (240 mg/dl) after week 6 to week 12, and >11.1 mmol/l (200 mg/dl) after week 12 to week 26.
Comparator	Placebo
Outcome measures	eGFR Creatinine level Adverse events Overall safety and tolerability were assessed by adverse event (AE) reports, safety laboratory tests, vital sign measurements, physical examinations and 12-lead electrocardiograms. Selected AEs of interest, including genital mycotic infections and urinary tract infections (UTIs), were prespecified for additional data collection. Events of hypoglycaemia were collected using a separate case report form that collected concurrent fingerstick glucose values and the presence of symptoms indicating a severe event (i.e. requiring the assistance of another individual or resulting in seizure or loss of consciousness). Urine Albumin Creatinine Ratio HbA1c

	he prespecified primary efficacy endpoint was the change from baseline in HbA1c at week 26. Prespecified secondary efficacy endpoints evaluated at week 26 were the proportion of subjects reaching HbA1c <7.0% and change from baseline in FPG. Blood pressure change from baseline in blood pressure (BP) Body weight percent change in baseline in body weight Fasting plasma lipids percent change in baseline fasting plasma lipids
Number of participants	272 randomised
Duration of follow-up	52 weeks
Loss to follow-up	35 discontinued treatment, 10 experienced an adverse event, 8 withdrew consent, 13 dropped out for other reasons, 2 had protocol violation, 1 was noncompliant, 1 died. However, all were analysed via intention to treat, except for 3 who did not initially take the study drug.
Methods of analysis	Efficacy analyses were conducted using the modified intent-to-treat (mITT) population, which consisted of all randomized subjects who received ≥1 dose of study drug, according to the randomized treatment assignment. The last observation carried forward (LOCF) approach was used to impute missing data. If subjects received rescue therapy, all post rescue data were censored and the last postbaseline value prior to the initiation of rescue therapy was used for analyses. Safety analyses were performed in randomized subjects who received ≥1dose of study drug according to the predominant treatment received.

Additional comments	Primary and continuous secondary efficacy endpoints were assessed using an analysis of covariance (ancova) model with treatment and stratification factors as fixed effects and corresponding baseline values and baseline eGFR as covariates. Least squares (LS) mean differences and two-sided 95% confidence intervals (CIs) were estimated based on this model for the comparison of each canagliflozin group versus placebo. The categorical secondary endpoint (proportion of subjects reaching HbA1c < 7.0%) was analyzed using a logistic model with treatment and stratification factors as fixed effects and baseline HbA1c and eGFR values as covariates.

Study arms

Placebo (N = 90)

oral, once daily

canagliflozin 100 mg (N = 90)

oral, once daily

canagliflozin 300 mg (N = 89)

oral, once daily

Characteristics

Study-level characteristics

Characteristic	Study (N =)
% Female Sample size	n = 106 ; % = 39.4
Mean age (SD) Mean (SD)	68.5 (8.3)
BMI Mean (SD)	33 (6.2)
White Sample size	n = 215 ; % = 79.9
Black or African American Sample size	n = 5 ; % = 1.9
Asian Sample size	n = 27 ; % = 10
Other Sample size	n = 22 ; % = 8.2
eGFR ml/min/1.73 m2 Mean (SD)	39.4 (6.9)

Characteristic	Study (N =)
Median ACR, µg/mg Median (IQR)	30

Arm-level characteristics

Characteristic	Placebo (N = 90)	canagliflozin 100 mg (N = 90)	canagliflozin 300 mg (N = 89)
Female Sample size	n = 33 ; % = 36.7	n = 32 ; % = 35.6	n = 41 ; % = 46.1
Age, years Mean (SD)	68.2 (8.4)	69.5 (8.2)	67.9 (8.2)
White Sample size	n = 78 ; % = 86.7	n = 71 ; % = 78.9	n = 66 ; % = 74.2
Black or African American Sample size	n = 0 ; % = 0	n = 3 ; % = 3.3	n = 2 ; % = 2.2
Asian Sample size	n = 7 ; % = 7.8	n = 9 ; % = 10	n = 11 ; % = 12.4
Other Sample size	n = 5 ; % = 5.6	n = 7 ; % = 7.8	n = 10 ; % = 11.2

Characteristic	Placebo (N = 90)	canagliflozin 100 mg (N = 90)	canagliflozin 300 mg (N = 89)
BMI Mean (SD)	33.1 (6.5)	32.4 (5.5)	33.4 (6.5)
eGFR ml/min/1.73 m2 Mean (SD)	40.1 (6.8)	39.7 (6.9)	38.5 (6.9)
Median ACR µg/mg Median (IQR)	31.3 (*empty data* to *empty data*)	23.7 (*empty data* to *empty data*)	30.1 (*empty data* to *empty data*)

Critical appraisal - GUT Cochrane Risk of Bias tool (RoB 2.0) Normal RCT

Section	Question	Answer
Domain 1: Bias arising from the randomisation process	1. 1. Was the allocation sequence random?	Yes
Domain 1: Bias arising from the randomisation process	1. 2. Was the allocation sequence concealed until participants were enrolled and assigned to interventions?	Yes
Domain 1: Bias arising from the randomisation process	1.3 Did baseline differences between intervention groups suggest a problem with the randomisation process?	No

Section	Question	Answer
Domain 1: Bias arising from the randomisation process	Risk of bias judgement for the randomisation process	Low
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.1. Were participants aware of their assigned intervention during the trial?	No
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	No
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?	No/Probably no
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.4. If Y/PY to 2.3: Were these deviations from intended intervention balanced between groups?	Not applicable
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.5 If N/PN/NI to 2.4: Were these deviations likely to have affected the outcome?	Not applicable
Domain 2a: Risk of bias due to deviations from the intended	2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?	Yes

Section	Question	Answer
interventions (effect of assignment to intervention)		
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized?	Not applicable
Domain 2a: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention)	Risk of bias for deviations from the intended interventions (effect of assignment to intervention)	Low
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.1. Were participants aware of their assigned intervention during the trial?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.3. If Y/PY/NI to 2.1 or 2.2: Were important co-interventions balanced across intervention groups?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.4. Could failures in implementing the intervention have affected the outcome?	*N/A.*

Section	Question	Answer
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.5. Did study participants adhere to the assigned intervention regimen?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	2.6. If N/PN/NI to 2.3 or 2.5 or Y/PY/NI to 2.4: Was an appropriate analysis used to estimate the effect of adhering to the intervention?	*N/A.*
Domain 2b: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)	Risk of bias judgement for deviations from the intended interventions (effect of adhering to intervention)	*N/A.*
Domain 3. Bias due to missing outcome data	3.1 Were data for this outcome available for all, or nearly all, participants randomised?	Probably yes
Domain 3. Bias due to missing outcome data	3.2 If N/PN/NI to 3.1: Is there evidence that result was not biased by missing outcome data?	Not applicable
Domain 3. Bias due to missing outcome data	3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	Not applicable
Domain 3. Bias due to missing outcome data	3.4 If Y/PY/NI to 3.3: Do the proportions of missing outcome data differ between intervention groups?	Not applicable
Domain 3. Bias due to missing outcome data	3.5 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	Not applicable

Section	Question	Answer
Domain 3. Bias due to missing outcome data	Risk-of-bias judgement for missing outcome data	Some concerns *(modified intention to treat was used with last observation carried forward for missing data - however it was not clear how much missing data there was or whether the amount differed between experimental arms.)*
Domain 4. Bias in measurement of the outcome	4.1 Was the method of measuring the outcome inappropriate?	Probably no
Domain 4. Bias in measurement of the outcome	4.2 Could measurement or ascertainment of the outcome have differed between intervention groups ?	Probably no
Domain 4. Bias in measurement of the outcome	4.3 If N/PN/NI to 4.1 and 4.2: Were outcome assessors aware of the intervention received by study participants ?	No
Domain 4. Bias in measurement of the outcome	4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?	Not applicable
Domain 4. Bias in measurement of the outcome	4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?	Not applicable
Domain 4. Bias in measurement of the outcome	Risk-of-bias judgement for measurement of the outcome	Low
Domain 5. Bias in selection of the reported result	5.1 Was the trial analysed in accordance with a pre-specified plan that was finalised before unblinded outcome data were available for analysis ?	Yes

Section	Question	Answer
Domain 5. Bias in selection of the reported result	5.2 Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	No/Probably no
Domain 5. Bias in selection of the reported result	5.3 Is the numerical result being assessed likely to have been selected, on the basis of the results, from multiple analyses of the data?	No/Probably no
Domain 5. Bias in selection of the reported result	Risk-of-bias judgement for selection of the reported result	Low
Overall bias and Directness	Risk of bias judgement	Low
Overall bias and Directness	Overall Directness	Directly applicable

Appendix E – Forest plots

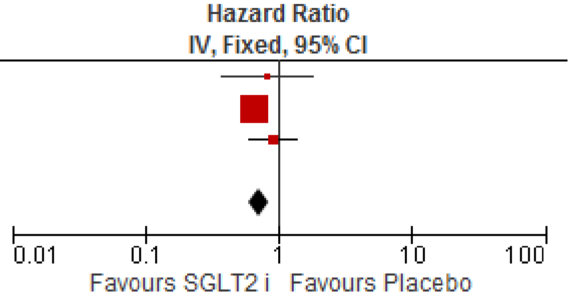
In all plots, eGFR is reported in units of ml/min/1.73 m2. The analysis is subgrouped by individual SGLT2 inhibitor only when substantial heterogeneity (I2>50%) was identified.

<u>SGLT2 vs Placebo primary analysis</u>

Renal composite – End stage kidney disease, doubling serum creatinine, renal death

Study or Subgroup	log[Hazard Ratio]	SE	SGLT-2 i Total	Placebo Total	Weight	Hazard Ratio IV, Fixed, 95% CI
CANVAS (eGFR 30-60) - canagliflozin	-0.211	0.399	1110	929	5.5%	0.81 [0.37, 1.77]
CREDENCE (eGFR 30-90)- canagliflozin	-0.416	0.108	2202	2199	75.7%	0.66 [0.53, 0.82]
VERTIS CV (eGFR 30-60) Ertugliflozin (1)	-0.105	0.217	1199	608	18.8%	0.90 [0.59, 1.38]
Total (95% CI)			**4511**	**3736**	**100.0%**	**0.71 [0.59, 0.85]**

Heterogeneity: Chi² = 1.77, df = 2 (P = 0.41); I² = 0%
Test for overall effect: Z = 3.69 (P = 0.0002)

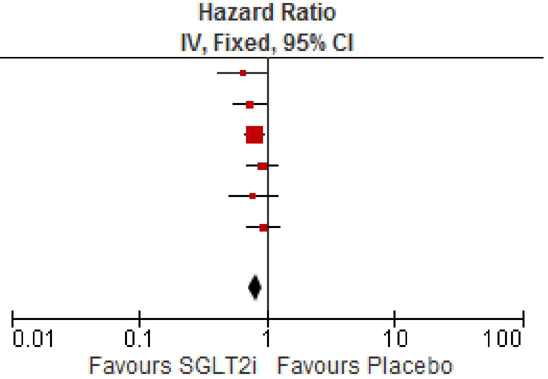


<u>Footnotes</u>
(1) Pooled data for 5 and 15mg doses

Cardiovascular composite: 3-point MACE

Study or Subgroup	log[Hazard Ratio]	SE	SGLT2i Total	Placebo Total	Weight	Hazard Ratio IV, Fixed, 95% CI
CANVAS (eGFR 30-45) - canagliflozin	-0.431	0.235	277	277	6.1%	0.65 [0.41, 1.03]
CANVAS (eGFR 45-60) - canagliflozin	-0.313	0.149	743	743	15.1%	0.73 [0.55, 0.98]
CREDENCE (eGFR 30-90)- canagliflozin	-0.223	0.089	2202	2199	42.3%	0.80 [0.67, 0.95]
DECLARE-TIMI (eGFR less than 60) - dapagliflozin	-0.083	0.147	606	659	15.5%	0.92 [0.69, 1.23]
EMPA-REG (eGFR 30-45) empagliflozin (1)	-0.236	0.222	381	189	6.8%	0.79 [0.51, 1.22]
EMPA-REG (eGFR 45-60) empagliflozin (2)	-0.072	0.154	831	418	14.1%	0.93 [0.69, 1.26]
Total (95% CI)			**5040**	**4485**	**100.0%**	**0.81 [0.73, 0.91]**

Heterogeneity: Chi² = 2.94, df = 5 (P = 0.71); I² = 0%
Test for overall effect: Z = 3.57 (P = 0.0004)

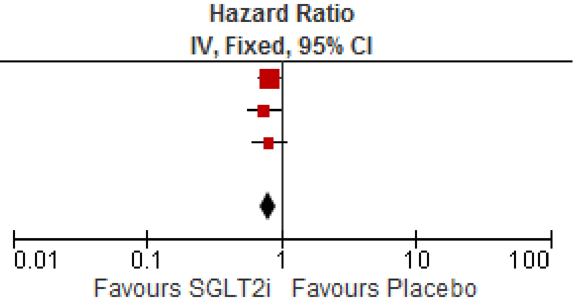


Footnotes
(1) Pooled 10 and 25mg doses
(2) Pooled 10 and 25mg doses

All-cause mortality

Study or Subgroup	log[Hazard Ratio]	SE	SGLT2i Total	Placebo Total	Weight	Hazard Ratio IV, Fixed, 95% CI
CREDENCE (eGFR 30-90)- canagliflozin	-0.186	0.103	2202	2199	51.0%	0.83 [0.68, 1.02]
DAPA-CKD (eGFR 25-75) - dapagliflozin	-0.301	0.143	1455	1451	26.5%	0.74 [0.56, 0.98]
EMPA-REG (allCKD) empagliflozin (1)	-0.211	0.155	757	752	22.5%	0.81 [0.60, 1.10]
Total (95% CI)			**4414**	**4402**	**100.0%**	**0.80 [0.69, 0.93]**

Heterogeneity: Chi² = 0.43, df = 2 (P = 0.81); I² = 0%
Test for overall effect: Z = 3.02 (P = 0.003)

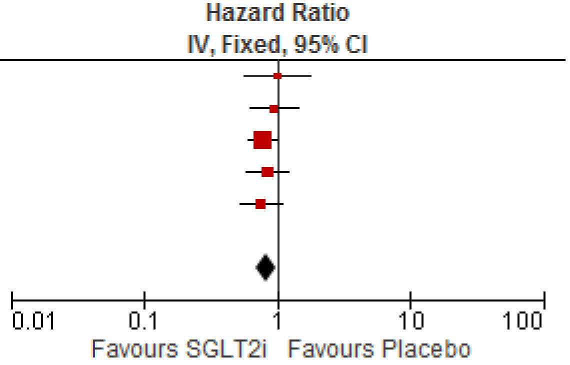


Footnotes
(1) 10 mg dose

Cardiovascular death

Study or Subgroup	log[Hazard Ratio]	SE	SGLT2 i Total	Placebo Total	Weight	Hazard Ratio IV, Fixed, 95% CI
CANVAS (eGFR 30-45) - canagliflozin	0.01	0.294	277	277	7.5%	1.01 [0.57, 1.80]
CANVAS (eGFR 45-60) - canagliflozin	-0.062	0.211	743	743	14.6%	0.94 [0.62, 1.42]
CREDENCE (eGFR 30-90)- canagliflozin	-0.248	0.126	2202	2199	40.8%	0.78 [0.61, 1.00]
DAPA-CKD (eGFR 25-75) - dapagliflozin	-0.163	0.183	1455	1451	19.4%	0.85 [0.59, 1.22]
EMPA-REG (allCKD) empagliflozin (1)	-0.274	0.191	757	752	17.8%	0.76 [0.52, 1.11]
Total (95% CI)			**5434**	**5422**	**100.0%**	**0.83 [0.71, 0.97]**

Heterogeneity: Chi² = 1.26, df = 4 (P = 0.87); I² = 0%
Test for overall effect: Z = 2.36 (P = 0.02)

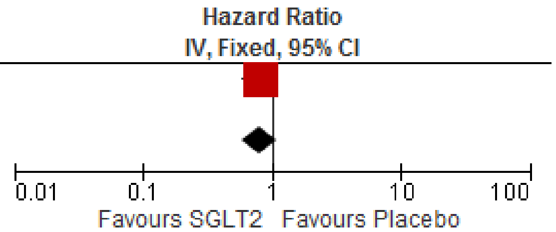


Footnotes
(1) 10 mg dose

Non-fatal MI

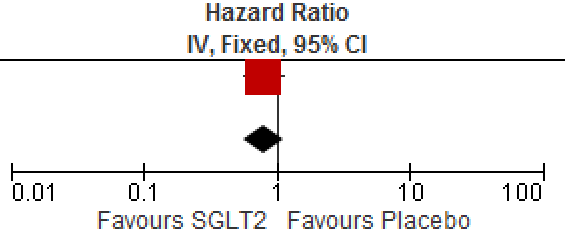
Study or Subgroup	log[Hazard Ratio]	SE	SGLT2 Total	Placebo Total	Weight	Hazard Ratio IV, Fixed, 95% CI
CREDENCE (eGFR 30-90)- canagliflozin	-0.21072103	0.158914	2202	2199	100.0%	0.81 [0.59, 1.11]
Total (95% CI)			**2202**	**2199**	**100.0%**	**0.81 [0.59, 1.11]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.33 (P = 0.18)



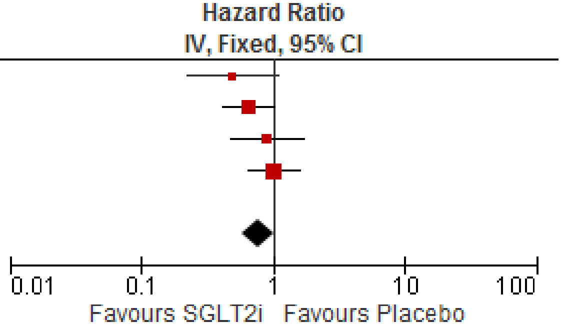
Non-fatal stroke

Study or Subgroup	log[Hazard Ratio]	SE	SGLT2 Total	Placebo Total	Weight	Hazard Ratio IV, Fixed, 95% CI
CREDENCE (eGFR 30-90)- canagliflozin	-0.22314	0.183566	2202	2199	100.0%	0.80 [0.56, 1.15]
Total (95% CI)			**2202**	**2199**	**100.0%**	**0.80 [0.56, 1.15]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.22 (P = 0.22)						



Fatal/non-fatal MI

Study or Subgroup	log[Hazard Ratio]	SE	SGLT2i Total	Placebo Total	Weight	Hazard Ratio IV, Fixed, 95% CI
CANVAS (eGFR 30-45) - canagliflozin	-0.712	0.404	277	277	11.8%	0.49 [0.22, 1.08]
CANVAS (eGFR 45-60) - canagliflozin	-0.431	0.237	743	743	34.3%	0.65 [0.41, 1.03]
EMPA-REG (eGFR 30-45) empagliflozin (1)	-0.105	0.32	381	189	18.8%	0.90 [0.48, 1.69]
EMPA-REG (eGFR 45-60) empagliflozin (2)	0.01	0.234	831	418	35.1%	1.01 [0.64, 1.60]
Total (95% CI)			**2232**	**1627**	**100.0%**	**0.78 [0.59, 1.02]**
Heterogeneity: Chi² = 3.33, df = 3 (P = 0.34); I² = 10%						
Test for overall effect: Z = 1.79 (P = 0.07)						



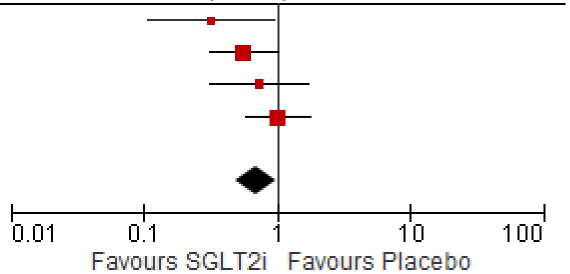
Footnotes
(1) Pooled 10 and 25mg doses
(2) Pooled 10 and 25mg doses

Fatal/non-fatal stroke

Study or Subgroup	log[Hazard Ratio]	SE	SGLT2i Total	Placebo Total	Weight	Hazard Ratio IV, Fixed, 95% CI
CANVAS (eGFR 30-45) - canagliflozin	-1.14	0.553	277	277	10.2%	0.32 [0.11, 0.95]
CANVAS (eGFR 45-60) - canagliflozin	-0.58	0.299	743	743	34.9%	0.56 [0.31, 1.01]
EMPA-REG (eGFR 30-45) empagliflozin	-0.301	0.43	381	189	16.9%	0.74 [0.32, 1.72]
EMPA-REG (eGFR 45-60) empagliflozin	0.02	0.286	831	418	38.1%	1.02 [0.58, 1.79]
Total (95% CI)			**2232**	**1627**	**100.0%**	**0.70 [0.49, 0.98]**

Heterogeneity: Chi² = 4.31, df = 3 (P = 0.23); I² = 30%
Test for overall effect: Z = 2.05 (P = 0.04)

Hazard Ratio
IV, Fixed, 95% CI

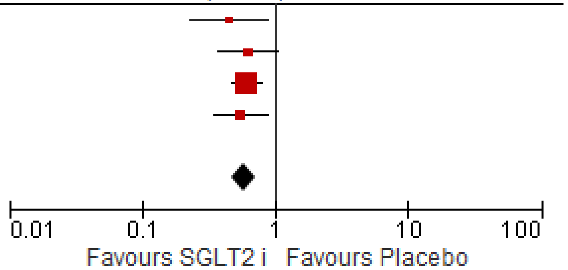


Hospitalisation for heart failure

Study or Subgroup	log[Hazard Ratio]	SE	SGLT2i Total	Placebo Total	Weight	Hazard Ratio IV, Fixed, 95% CI
CANVAS (eGFR 30-45) - canagliflozin	-0.799	0.342	277	277	8.9%	0.45 [0.23, 0.88]
CANVAS (eGFR 45-60) - canagliflozin	-0.478	0.261	743	743	15.2%	0.62 [0.37, 1.03]
CREDENCE (eGFR 30-90)- canagliflozin	-0.494	0.136	2202	2199	56.1%	0.61 [0.47, 0.80]
EMPA-REG (allCKD) empagliflozin (1)	-0.598	0.229	757	752	19.8%	0.55 [0.35, 0.86]
Total (95% CI)			**3979**	**3971**	**100.0%**	**0.58 [0.48, 0.71]**

Heterogeneity: Chi² = 0.81, df = 3 (P = 0.85); I² = 0%
Test for overall effect: Z = 5.29 (P < 0.00001)

Hazard Ratio
IV, Fixed, 95% CI

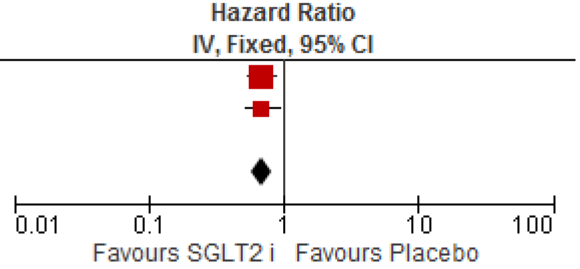


Footnotes
(1) 10 mg dose

End stage kidney disease

Study or Subgroup	log[Hazard Ratio]	SE	SGLT-2 i Total	Placebo Total	Weight	Hazard Ratio IV, Fixed, 95% CI
CREDENCE (eGFR 30-90)- canagliflozin	-0.386	0.119	2202	2199	61.7%	0.68 [0.54, 0.86]
DAPA-CKD (eGFR 25-75) - dapagliflozin	-0.371	0.151	1455	1451	38.3%	0.69 [0.51, 0.93]
Total (95% CI)			**3657**	**3650**	**100.0%**	**0.68 [0.57, 0.82]**

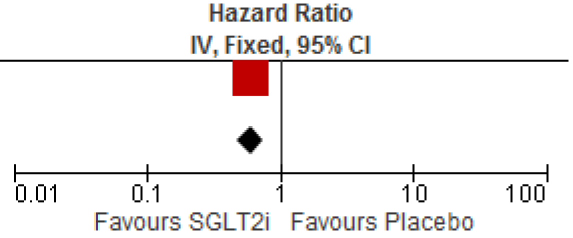
Heterogeneity: Chi² = 0.01, df = 1 (P = 0.94); I² = 0%
Test for overall effect: Z = 4.07 (P < 0.0001)



Doubling of serum creatinine

Study or Subgroup	log[Hazard Ratio]	SE	SGLT2i Total	Placebo Total	Weight	Hazard Ratio IV, Fixed, 95% CI
CREDENCE (eGFR 30-90)- canagliflozin	-0.511	0.117	2202	2199	100.0%	0.60 [0.48, 0.75]
Total (95% CI)			**2202**	**2199**	**100.0%**	**0.60 [0.48, 0.75]**

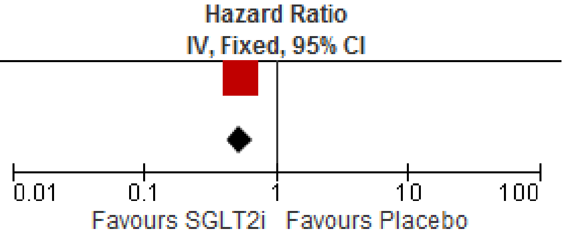
Heterogeneity: Not applicable
Test for overall effect: Z = 4.37 (P < 0.0001)



eGFR reduction >50%

Study or Subgroup	log[Hazard Ratio]	SE	SGLT2i Total	Placebo Total	Weight	Hazard Ratio IV, Fixed, 95% CI
DAPA-CKD (eGFR 25-75) - dapagliflozin	-0.63	0.119	1455	1451	100.0%	0.53 [0.42, 0.67]
Total (95% CI)			**1455**	**1451**	**100.0%**	**0.53 [0.42, 0.67]**

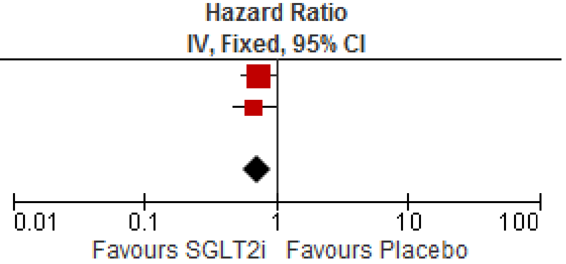
Heterogeneity: Not applicable
Test for overall effect: Z = 5.29 (P < 0.00001)



Dialysis

Study or Subgroup	log[Hazard Ratio]	SE	SGLT2i Total	Placebo Total	Weight	Hazard Ratio IV, Fixed, 95% CI
CREDENCE (eGFR 30-90)- canagliflozin (1)	-0.301	0.153	2202	2199	59.9%	0.74 [0.55, 1.00]
DAPA-CKD (eGFR 25-75) - dapagliflozin	-0.386	0.187	1455	1454	40.1%	0.68 [0.47, 0.98]
Total (95% CI)			**3657**	**3653**	**100.0%**	**0.72 [0.57, 0.90]**

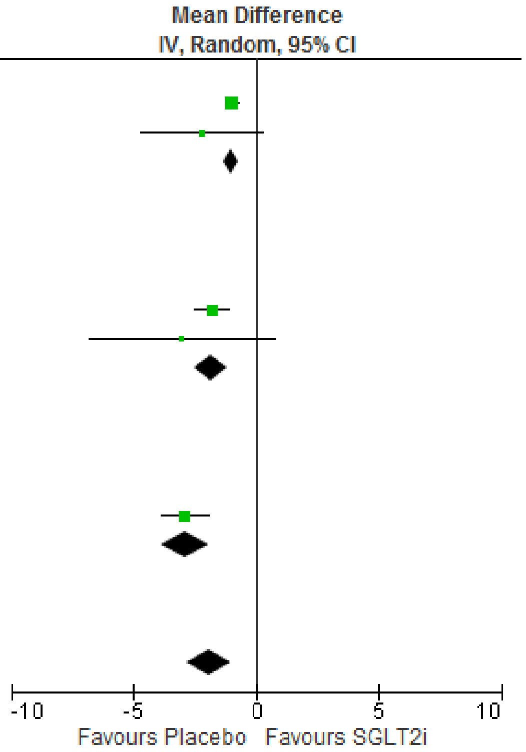
Heterogeneity: Chi² = 0.12, df = 1 (P = 0.72); I² = 0%
Test for overall effect: Z = 2.83 (P = 0.005)



Footnotes
(1) Only report dialysis + transplantation, but transplantation expected to contribute few events

eGFR 6 months (mean difference – in ml/min/1.73 m^{2})

Study or Subgroup	SGLT2i Mean	SGLT2i SD	SGLT2i Total	Placebo Mean	Placebo SD	Placebo Total	Weight	Mean Difference IV, Random, 95% CI
1.12.1 Canagliflozin								
CREDENCE (eGFR 30-90)- canagliflozin (1)	-4.8	4.48	2005	-3.8	4.45	1985	32.6%	-1.00 [-1.28, -0.72]
Yale 2013/14 (eGFR 30-50) Canagliflozin (2)	-3.6	8.5381	90	-1.4	8.5381	90	9.7%	-2.20 [-4.69, 0.29]
Subtotal (95% CI)			**2095**			**2075**	**42.3%**	**-1.01 [-1.29, -0.74]**
Heterogeneity: Tau² = 0.00; Chi² = 0.88, df = 1 (P = 0.35); I² = 0%								
Test for overall effect: Z = 7.22 (P < 0.00001)								
1.12.2 Dapagliflozin								
DECLARE-TIMI (eGFR less than 60) - dapagliflozin (3)	0.2	6.1097	576	2	6.3243	617	28.1%	-1.80 [-2.51, -1.09]
DERIVE (eGFR 45 to 60) - dapagliflozin (4)	-3.9	18.0831	142	-0.9	14.3625	140	5.0%	-3.00 [-6.81, 0.81]
Subtotal (95% CI)			**718**			**757**	**33.0%**	**-1.84 [-2.53, -1.15]**
Heterogeneity: Tau² = 0.00; Chi² = 0.37, df = 1 (P = 0.54); I² = 0%								
Test for overall effect: Z = 5.20 (P < 0.00001)								
1.12.3 Ertugliflozin								
VERTIS RENAL (eGFR 30-60) Ertugliflozin (5)	-2.6	4.84	313	0.3	5.0252	154	24.6%	-2.90 [-3.86, -1.94]
Subtotal (95% CI)			**313**			**154**	**24.6%**	**-2.90 [-3.86, -1.94]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 5.93 (P < 0.00001)								
Total (95% CI)			**3126**			**2986**	**100.0%**	**-1.91 [-2.83, -0.99]**
Heterogeneity: Tau² = 0.65; Chi² = 18.23, df = 4 (P = 0.001); I² = 78%								
Test for overall effect: Z = 4.07 (P < 0.0001)								
Test for subgroup differences: Chi² = 16.98, df = 2 (P = 0.0002), I² = 88.2%								



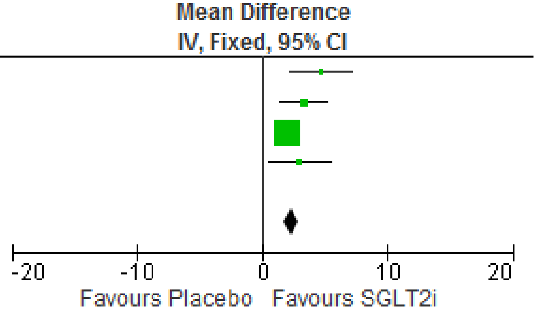
Footnotes
(1) Reported as change from baseline
(2) Used 100mg dose. Read from Graph. Reported as change from baseline.
(3) Read from graph. Reported as change from baseline
(4) Read from graph. 24 weeks follow up. Reported as change from baseline.
(5) Combined 5 and 15mg doses. Read from Graph. Reported as change from baseline.

(6) eGFR 6 months (ml/min/1.73m2, higher values are better)

eGFR last available data point > 2years

Study or Subgroup	SGLT2i Mean	SGLT2i SD	SGLT2i Total	Placebo Mean	Placebo SD	Placebo Total	Weight	Mean Difference IV, Fixed, 95% CI
CANVAS (eGFR 30-45) - canagliflozin (1)	40.5	4.9928	45	35.8	4.2789	17	6.1%	4.70 [2.20, 7.20]
CANVAS (eGFR 45-60) - canagliflozin (2)	53.8	7.8305	166	50.5	6.9963	86	10.6%	3.30 [1.40, 5.20]
DECLARE-TIMI (eGFR less than 60) - dapagliflozin (3)	-2	4.9702	382	-4	5.0288	391	77.0%	2.00 [1.30, 2.70]
VERTIS CV (eGFR 30-60) Ertugliflozin (4)	51	10.6755	197	48	9.384	87	6.3%	3.00 [0.53, 5.47]
Total (95% CI)			**790**			**581**	**100.0%**	**2.37 [1.75, 2.98]**

Heterogeneity: Chi² = 5.56, df = 3 (P = 0.14); I² = 46%
Test for overall effect: Z = 7.49 (P < 0.00001)



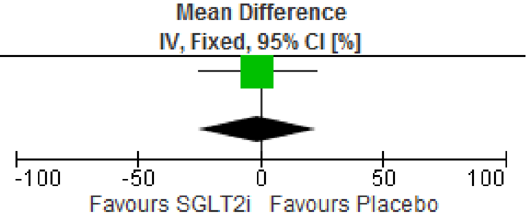
Footnotes
(1) Read from graph. Follow up 6 years
(2) Read from graph. Follow up 6 years
(3) Read from graph. Follow up 4 years. Reported as change from baseline
(4) Read from graph. Follow up 5 years

5) eGFR >2years (ml/min/1.73m2, higher values are better)

Percentage change from baseline UACR 6 months

Study or Subgroup	SGLT2i Mean [%]	SGLT2i SD [%]	SGLT2i Total	Placebo Mean [%]	Placebo SD [%]	Placebo Total	Weight	Mean Difference IV, Fixed, 95% CI [%]
DELIGHT (eGFR 20-80) - dapaglaflozin (1)	-2	104.6506	148	-1	103.2027	144	100.0%	-1.00 [-24.84, 22.84]
Total (95% CI)			**148**			**144**	**100.0%**	**-1.00 [-24.84, 22.84]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.08 (P = 0.93)



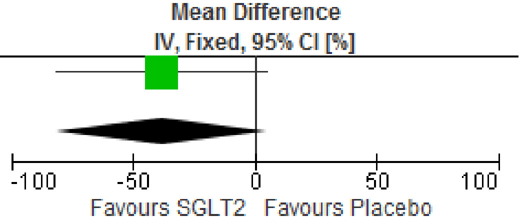
Footnotes
(1) Read from graph. 24 weeks follow up.

2) Percentage change from baseline uACR (%, lower values are better)

Percentage change from baseline UACR last available data point >2years

Study or Subgroup	SGLT2i Mean [%]	SGLT2i SD [%]	SGLT2i Total	Placebo Mean [%]	Placebo SD [%]	Placebo Total	Weight	Mean Difference IV, Fixed, 95% CI [%]
VERTIS CV (eGFR 30-60) Ertugliflozin (1)	32	101.268	276	70	227.3607	115	100.0%	-38.00 [-81.24, 5.24]
Total (95% CI)			**276**			**115**	**100.0%**	**-38.00 [-81.24, 5.24]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.72 (P = 0.08)



Footnotes
(1) Read from graph. 2.5 years follow up.

2)

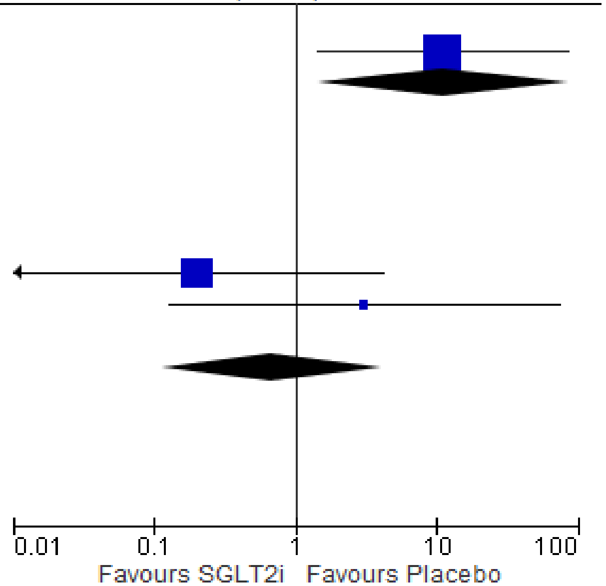
2) Percentage change from baseline uACR (%, lower values are better)

Diabetic Ketoacidosis SGLT2 subgroups (fixed effects model)

A fixed and random effects model was produced because there was substantial heterogeneity in the overall analysis (I^2>50%) which prompted an analysis by individual drug. However, there was no substantial heterogeneity in the subgroup estimates and so a fixed effects model was used for the subgroup estimates in the GRADE profiles, and for decision making. Both are shown here for completion.

	SGLT2i		Placebo			Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% CI
1.16.1 Canagliflozin						
CREDENCE (eGFR 30-90)- canagliflozin	11	2200	1	2197	100.0%	10.98 [1.42, 85.01]
Subtotal (95% CI)		**2200**		**2197**	**100.0%**	**10.98 [1.42, 85.01]**
Total events	11		1			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.30 (P = 0.02)						
1.16.2 Dapagliflozin						
DAPA-CKD (eGFR 25-75) - dapagliflozin	0	2149	2	2149	83.5%	0.20 [0.01, 4.16]
DELIGHT (eGFR 20-80) - dapaglaflozin	1	145	0	148	16.5%	3.06 [0.13, 74.55]
DERIVE (eGFR 45 to 60) - dapagliflozin	0	161	0	160		Not estimable
Subtotal (95% CI)		**2455**		**2457**	**100.0%**	**0.67 [0.11, 4.00]**
Total events	1		2			
Heterogeneity: Chi² = 1.48, df = 1 (P = 0.22); I² = 32%						
Test for overall effect: Z = 0.44 (P = 0.66)						

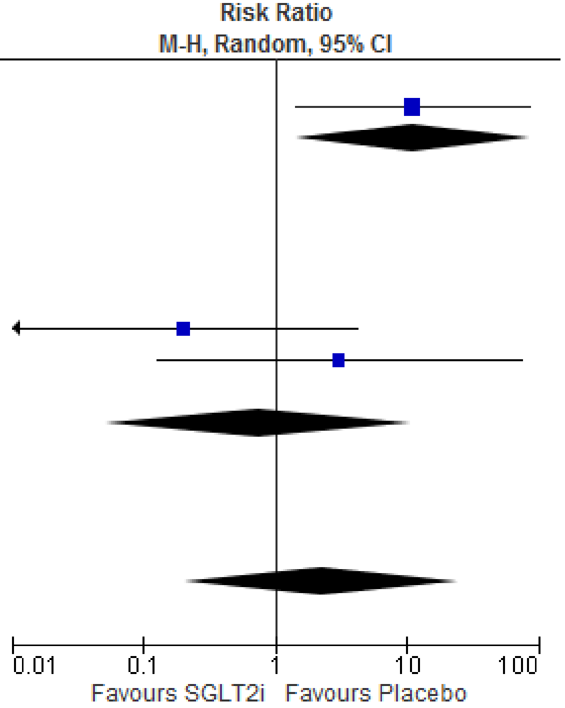
Risk Ratio
M-H, Fixed, 95% CI



Test for subgroup differences: Chi² = 4.07, df = 1 (P = 0.04), I² = 75.4%

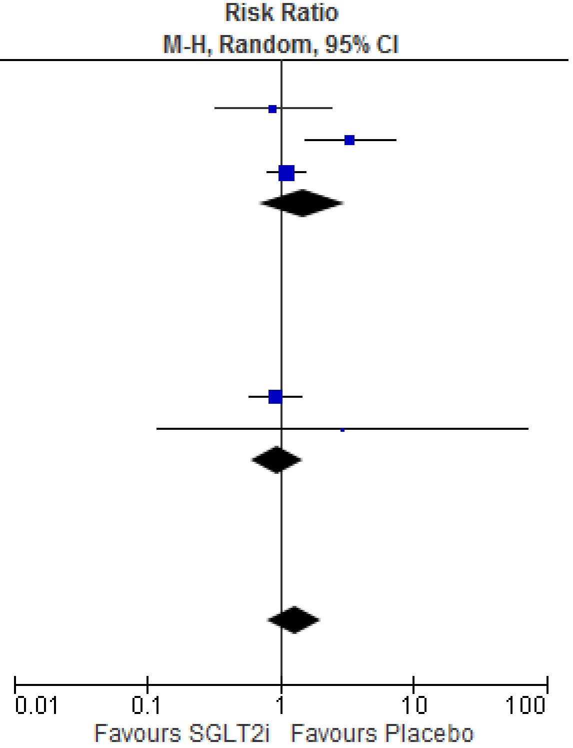
Diabetic Ketoacidosis with total (random effects model)

Study or Subgroup	SGLT2i Events	SGLT2i Total	Placebo Events	Placebo Total	Weight	Risk Ratio M-H, Random, 95% CI
1.16.1 Canagliflozin						
CREDENCE (eGFR 30-90)- canagliflozin	11	2200	1	2197	41.1%	10.98 [1.42, 85.01]
Subtotal (95% CI)		**2200**		**2197**	**41.1%**	**10.98 [1.42, 85.01]**
Total events	11		1			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.30 (P = 0.02)						
1.16.2 Dapagliflozin						
DAPA-CKD (eGFR 25-75) - dapagliflozin	0	2149	2	2149	30.2%	0.20 [0.01, 4.16]
DELIGHT (eGFR 20-80) - dapaglaflozin	1	145	0	148	28.7%	3.06 [0.13, 74.55]
DERIVE (eGFR 45 to 60) - dapagliflozin	0	161	0	160		Not estimable
Subtotal (95% CI)		**2455**		**2457**	**58.9%**	**0.75 [0.05, 10.86]**
Total events	1		2			
Heterogeneity: Tau² = 1.21; Chi² = 1.48, df = 1 (P = 0.22); I² = 32%						
Test for overall effect: Z = 0.21 (P = 0.83)						
Total (95% CI)		**4655**		**4654**	**100.0%**	**2.27 [0.21, 24.73]**
Total events	12		3			
Heterogeneity: Tau² = 2.52; Chi² = 4.60, df = 2 (P = 0.10); I² = 57%						
Test for overall effect: Z = 0.67 (P = 0.50)						
Test for subgroup differences: Chi² = 2.44, df = 1 (P = 0.12), I² = 59.1%						



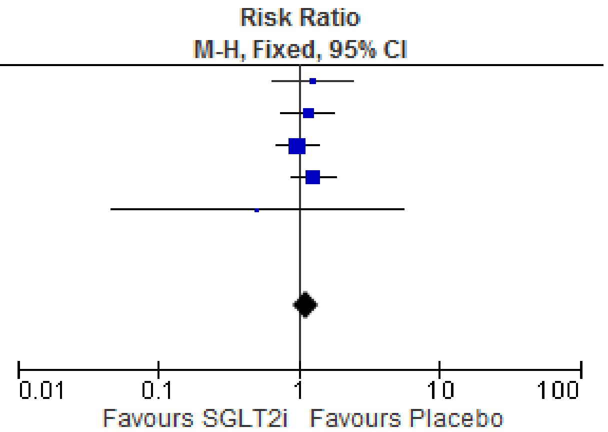
Amputation

Study or Subgroup	SGLT2i Events	SGLT2i Total	Placebo Events	Placebo Total	Weight	Risk Ratio M-H, Random, 95% CI
1.17.1 Canagliflozin						
CANVAS (eGFR 30-45) - canagliflozin	7	277	8	277	14.1%	0.88 [0.32, 2.38]
CANVAS (eGFR 45-60) - canagliflozin	27	743	8	743	19.0%	3.38 [1.54, 7.38]
CREDENCE (eGFR 30-90)- canagliflozin	70	2200	63	2197	34.9%	1.11 [0.79, 1.55]
Subtotal (95% CI)		**3220**		**3217**	**67.9%**	**1.48 [0.70, 3.13]**
Total events	104		79			
Heterogeneity: Tau² = 0.31; Chi² = 7.20, df = 2 (P = 0.03); I² = 72%						
Test for overall effect: Z = 1.02 (P = 0.31)						
1.17.2 Dapagliflozin						
DAPA-CKD (eGFR 25-75) - dapagliflozin	35	2149	38	2149	30.1%	0.92 [0.58, 1.45]
DELIGHT (eGFR 20-80) - dapaglaflozin	1	148	0	145	2.0%	2.94 [0.12, 71.57]
Subtotal (95% CI)		**2297**		**2294**	**32.1%**	**0.94 [0.60, 1.48]**
Total events	36		38			
Heterogeneity: Tau² = 0.00; Chi² = 0.50, df = 1 (P = 0.48); I² = 0%						
Test for overall effect: Z = 0.26 (P = 0.80)						
Total (95% CI)		**5517**		**5511**	**100.0%**	**1.28 [0.81, 2.02]**
Total events	140		117			
Heterogeneity: Tau² = 0.13; Chi² = 8.89, df = 4 (P = 0.06); I² = 55%						
Test for overall effect: Z = 1.05 (P = 0.29)						
Test for subgroup differences: Chi² = 1.01, df = 1 (P = 0.31), I² = 1.0%						



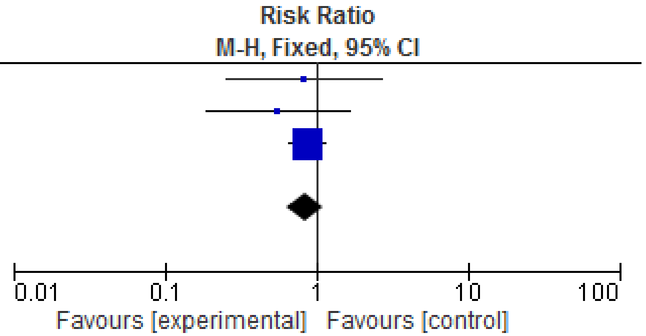
Fracture

Study or Subgroup	SGLT2i Events	SGLT2i Total	Placebo Events	Placebo Total	Weight	Risk Ratio M-H, Fixed, 95% CI
CANVAS (eGFR 30-45) - canagliflozin	19	277	15	277	8.7%	1.27 [0.66, 2.44]
CANVAS (eGFR 45-60) - canagliflozin	42	743	36	743	20.9%	1.17 [0.76, 1.80]
CREDENCE (eGFR 30-90)- canagliflozin	67	2200	68	2197	39.6%	0.98 [0.71, 1.37]
DAPA-CKD (eGFR 25-75) - dapagliflozin	65	2149	51	2149	29.6%	1.27 [0.89, 1.83]
DELIGHT (eGFR 20-80) - dapaglaflozin	1	145	2	148	1.2%	0.51 [0.05, 5.57]
DERIVE (eGFR 45 to 60) - dapagliflozin	0	160	0	161		Not estimable
Total (95% CI)		**5674**		**5675**	**100.0%**	**1.13 [0.92, 1.38]**
Total events	194		172			
Heterogeneity: Chi² = 1.65, df = 4 (P = 0.80); I² = 0%						
Test for overall effect: Z = 1.17 (P = 0.24)						



Acute kidney injury

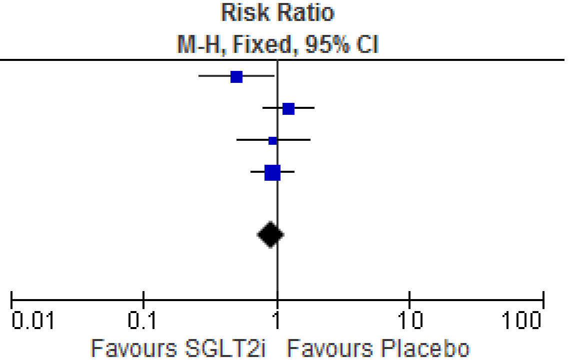
Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk Ratio M-H, Fixed, 95% CI
CANVAS (eGFR 30-45) - canagliflozin	5	277	6	277	5.3%	0.83 [0.26, 2.70]
CANVAS (eGFR 45-60) - canagliflozin	5	743	9	743	8.0%	0.56 [0.19, 1.65]
CREDENCE (eGFR 30-90)- canagliflozin	86	2200	98	2197	86.7%	0.88 [0.66, 1.16]
Total (95% CI)		**3220**		**3217**	**100.0%**	**0.85 [0.65, 1.11]**
Total events	96		113			
Heterogeneity: Chi² = 0.63, df = 2 (P = 0.73); I² = 0%						
Test for overall effect: Z = 1.21 (P = 0.23)						



Hypoglycaemia

Study or Subgroup	SGLT2i Events	SGLT2i Total	Placebo Events	Placebo Total	Weight	Risk Ratio M-H, Fixed, 95% CI
DAPA-CKD (eGFR 25-75) - dapagliflozin (1)	14	2149	28	2149	23.0%	0.50 [0.26, 0.95]
DELIGHT (eGFR 20-80) - dapaglaflozin (2)	35	145	29	148	23.6%	1.23 [0.80, 1.90]
DERIVE (eGFR 45 to 60) - dapagliflozin (3)	17	160	18	161	14.8%	0.95 [0.51, 1.78]
VERTIS RENAL (eGFR 30-60) Ertugliflozin (4)	67	313	35	154	38.6%	0.94 [0.66, 1.35]
Total (95% CI)		**2767**		**2612**	**100.0%**	**0.91 [0.72, 1.15]**
Total events	133		110			

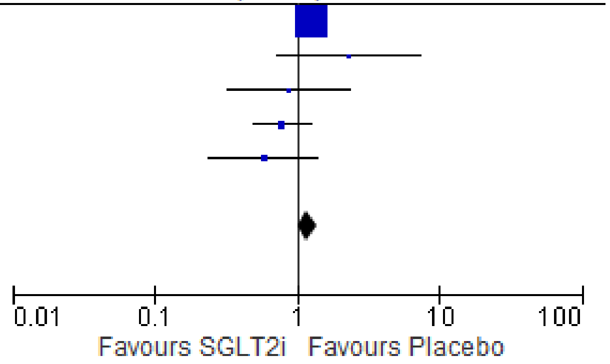
Heterogeneity: Chi² = 5.29, df = 3 (P = 0.15); I² = 43%
Test for overall effect: Z = 0.79 (P = 0.43)



Footnotes
(1) Serious hypoglycaemia
(2) Combined events from major and minor hypoglycaemia (other glycaemia may double count events)
(3) Used corrected data from erratum (2019)
(4) Symptomatic hypoglycaemia, week 52. Combined 5 and 15mg doses.

Genitourinary infection

Study or Subgroup	SGLT2i Events	SGLT2i Total	Placebo Events	Placebo Total	Weight	Risk Ratio M-H, Fixed, 95% CI	Risk Ratio M-H, Fixed, 95% CI
CREDENCE (eGFR 30-90)- canagliflozin (1)	295	2200	234	2197	80.7%	1.26 [1.07, 1.48]	
DELIGHT (eGFR 20-80) - dapaglaflozin (2)	9	145	4	148	1.4%	2.30 [0.72, 7.29]	
DERIVE (eGFR 45 to 60) - dapagliflozin (3)	7	160	8	161	2.7%	0.88 [0.33, 2.37]	
VERTIS RENAL (eGFR 30-60) Ertugliflozin (4)	38	313	24	154	11.1%	0.78 [0.49, 1.25]	
Yale 2013/14 (eGFR 30-50) Canagliflozin (5)	7	90	12	90	4.1%	0.58 [0.24, 1.41]	
Total (95% CI)		**2908**		**2750**	**100.0%**	**1.18 [1.02, 1.37]**	
Total events	356		282				
Heterogeneity: $Chi^2 = 7.63$, $df = 4$ ($P = 0.11$); $I^2 = 48\%$							
Test for overall effect: $Z = 2.23$ ($P = 0.03$)							

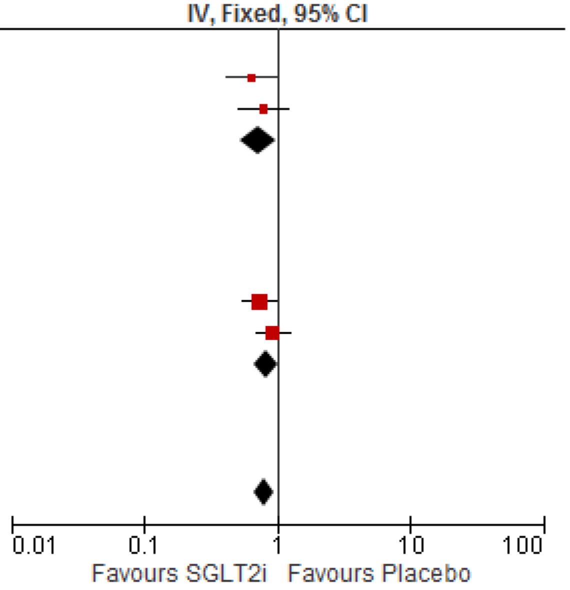


Footnotes
(1) Combined events for genital and urinary infections
(2) Combines events for urinary tract infection and genital infection
(3) Combined events for urinary tract infection and genital infection.
(4) Combined 5 and 15 mg doses. Combined events for urinary tract infection and genital infection. Week 52.
(5) Combined events for urinary tract infection and genital infection.

SGLT2 vs Placebo eGFR stratification

Cardiovascular composite: 3-point MACE

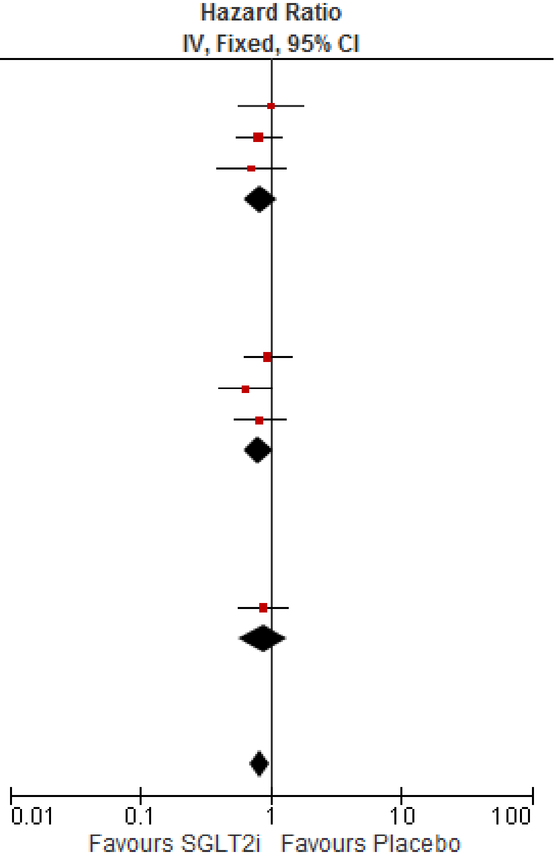
Study or Subgroup	log[]	SE	SGLT2i Total	Placebo Total	Weight	IV, Fixed, 95% CI
2.2.1 eGFR 30-45						
CANVAS (eGFR 30-45) - canagliflozin	-0.431	0.235	277	277	14.4%	0.65 [0.41, 1.03]
EMPA-REG (eGFR 30-45) empagliflozin (1)	-0.236	0.222	381	189	16.2%	0.79 [0.51, 1.22]
Subtotal (95% CI)			**658**	**466**	**30.6%**	**0.72 [0.53, 0.99]**
Heterogeneity: Chi² = 0.36, df = 1 (P = 0.55); I² = 0%						
Test for overall effect: Z = 2.03 (P = 0.04)						
2.2.2 eGFR 45-60						
CANVAS (eGFR 45-60) - canagliflozin	-0.313	0.149	743	743	35.9%	0.73 [0.55, 0.98]
EMPA-REG (eGFR 45-60) empagliflozin (2)	-0.072	0.154	831	418	33.6%	0.93 [0.69, 1.26]
Subtotal (95% CI)			**1574**	**1161**	**69.4%**	**0.82 [0.67, 1.01]**
Heterogeneity: Chi² = 1.26, df = 1 (P = 0.26); I² = 21%						
Test for overall effect: Z = 1.83 (P = 0.07)						
Total (95% CI)			**2232**	**1627**	**100.0%**	**0.79 [0.66, 0.94]**
Heterogeneity: Chi² = 2.09, df = 3 (P = 0.55); I² = 0%						
Test for overall effect: Z = 2.65 (P = 0.008)						
Test for subgroup differences: Chi² = 0.46, df = 1 (P = 0.50), I² = 0%						



Footnotes
(1) Pooled 10 and 25mg doses
(2) Combined 10 and 25mg doses

Cardiovascular death

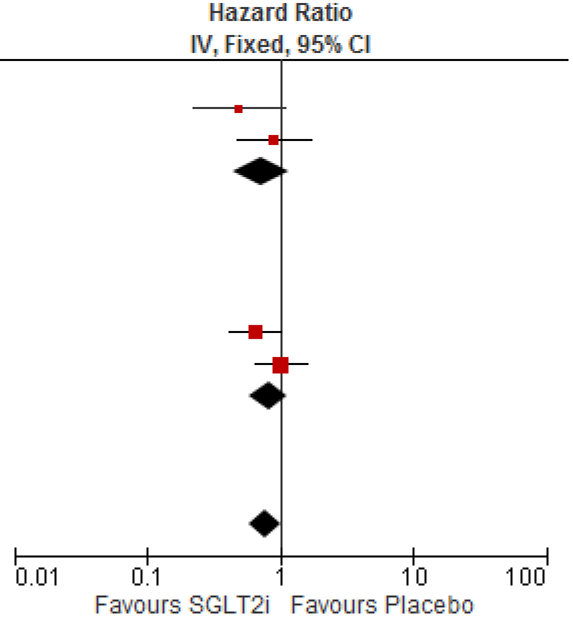
Study or Subgroup	log[Hazard Ratio]	SE	SGLT2 i Total	Placebo Total	Weight	Hazard Ratio IV, Fixed, 95% CI
2.4.1 eGFR (30-45)						
CANVAS (eGFR 30-45) - canagliflozin	0.01	0.294	277	277	9.3%	1.01 [0.57, 1.80]
CREDENCE (eGFR 30-45)- canagliflozin	-0.21072	0.205818	657	656	19.0%	0.81 [0.54, 1.21]
EMPA-REG (eGFR 30-45) empagliflozin (1)	-0.342	0.307	381	189	8.6%	0.71 [0.39, 1.30]
Subtotal (95% CI)			**1315**	**1122**	**36.9%**	**0.83 [0.62, 1.11]**
Heterogeneity: Chi² = 0.72, df = 2 (P = 0.70); I² = 0%						
Test for overall effect: Z = 1.25 (P = 0.21)						
2.4.2 eGFR (45-60)						
CANVAS (eGFR 45-60) - canagliflozin	-0.062	0.211	743	743	18.1%	0.94 [0.62, 1.42]
CREDENCE (eGFR 45-60)- canagliflozin	-0.44629	0.241288	640	639	13.9%	0.64 [0.40, 1.03]
EMPA-REG (eGFR 45-60) empagliflozin (2)	-0.198	0.234	831	418	14.7%	0.82 [0.52, 1.30]
Subtotal (95% CI)			**2214**	**1800**	**46.7%**	**0.80 [0.62, 1.04]**
Heterogeneity: Chi² = 1.45, df = 2 (P = 0.48); I² = 0%						
Test for overall effect: Z = 1.67 (P = 0.10)						
2.4.3 eGFR (60-90)						
CREDENCE (eGFR 60-90)- canagliflozin	-0.12783	0.221838	905	904	16.4%	0.88 [0.57, 1.36]
Subtotal (95% CI)			**905**	**904**	**16.4%**	**0.88 [0.57, 1.36]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.58 (P = 0.56)						
Total (95% CI)			**4434**	**3826**	**100.0%**	**0.83 [0.69, 0.98]**
Heterogeneity: Chi² = 2.29, df = 6 (P = 0.89); I² = 0%						
Test for overall effect: Z = 2.13 (P = 0.03)						
Test for subgroup differences: Chi² = 0.13, df = 2 (P = 0.94), I² = 0%						



Footnotes
(1) Pooled 10 and 25mg doses
(2) Pooled 10 and 25mg doses

Fatal/non-fatal MI

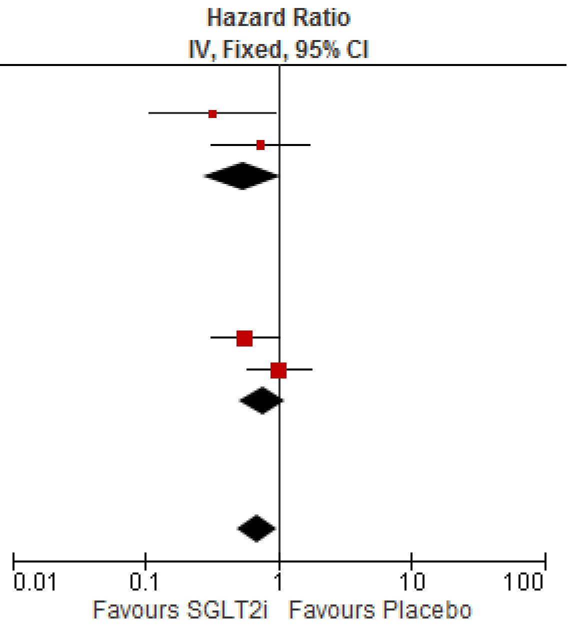
Study or Subgroup	log[Hazard Ratio]	SE	SGLT2i Total	Placebo Total	Weight	Hazard Ratio IV, Fixed, 95% CI
2.5.1 eGFR (30-45)						
CANVAS (eGFR 30-45) - canagliflozin	-0.712	0.404	277	277	11.8%	0.49 [0.22, 1.08]
EMPA-REG (eGFR 30-45) empagliflozin (1)	-0.105	0.32	381	189	18.8%	0.90 [0.48, 1.69]
Subtotal (95% CI)			**658**	**466**	**30.6%**	**0.71 [0.44, 1.16]**
Heterogeneity: Chi² = 1.39, df = 1 (P = 0.24); I² = 28%						
Test for overall effect: Z = 1.35 (P = 0.18)						
2.5.2 eGFR (45-60)						
CANVAS (eGFR 45-60) - canagliflozin	-0.431	0.237	743	743	34.3%	0.65 [0.41, 1.03]
EMPA-REG (eGFR 45-60) empagliflozin (2)	0.01	0.234	831	418	35.1%	1.01 [0.64, 1.60]
Subtotal (95% CI)			**1574**	**1161**	**69.4%**	**0.81 [0.59, 1.13]**
Heterogeneity: Chi² = 1.75, df = 1 (P = 0.19); I² = 43%						
Test for overall effect: Z = 1.25 (P = 0.21)						
Total (95% CI)			**2232**	**1627**	**100.0%**	**0.78 [0.59, 1.02]**
Heterogeneity: Chi² = 3.33, df = 3 (P = 0.34); I² = 10%						
Test for overall effect: Z = 1.79 (P = 0.07)						
Test for subgroup differences: Chi² = 0.19, df = 1 (P = 0.66), I² = 0%						



Footnotes
(1) Pooled 10 and 25mg doses
(2) Pooled 10 and 25mg doses

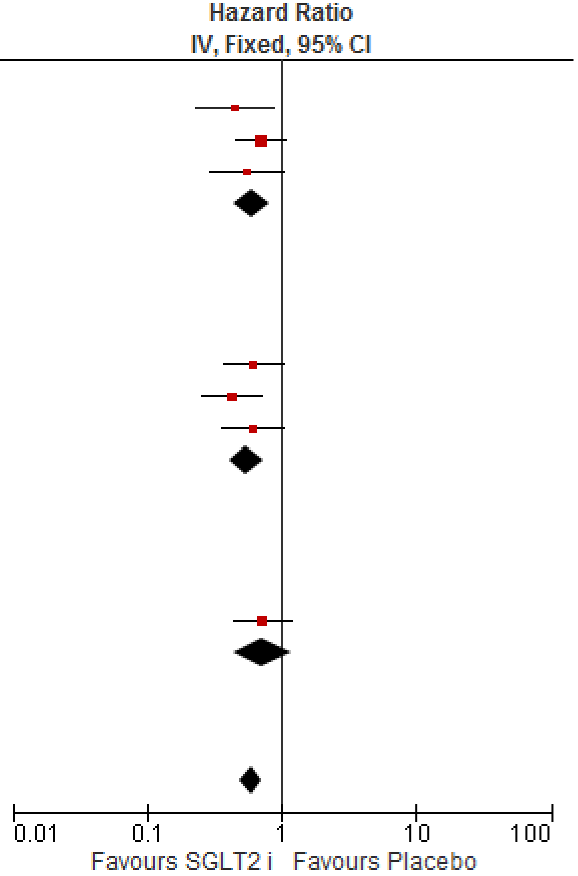
Fatal/non-fatal stroke

Study or Subgroup	log[Hazard Ratio]	SE	SGLT2i Total	Placebo Total	Weight	Hazard Ratio IV, Fixed, 95% CI
2.6.1 eGFR (30-45)						
CANVAS (eGFR 30-45) - canagliflozin	-1.14	0.553	277	277	10.2%	0.32 [0.11, 0.95]
EMPA-REG (eGFR 30-45) empagliflozin	-0.301	0.43	381	189	16.9%	0.74 [0.32, 1.72]
Subtotal (95% CI)			**658**	**466**	**27.0%**	**0.54 [0.28, 1.05]**
Heterogeneity: Chi² = 1.43, df = 1 (P = 0.23); I² = 30%						
Test for overall effect: Z = 1.82 (P = 0.07)						
2.6.2 eGFR 45-60						
CANVAS (eGFR 45-60) - canagliflozin	-0.58	0.299	743	743	34.9%	0.56 [0.31, 1.01]
EMPA-REG (eGFR 45-60) empagliflozin	0.02	0.286	831	418	38.1%	1.02 [0.58, 1.79]
Subtotal (95% CI)			**1574**	**1161**	**73.0%**	**0.77 [0.51, 1.15]**
Heterogeneity: Chi² = 2.10, df = 1 (P = 0.15); I² = 52%						
Test for overall effect: Z = 1.29 (P = 0.20)						
Total (95% CI)			**2232**	**1627**	**100.0%**	**0.70 [0.49, 0.98]**
Heterogeneity: Chi² = 4.31, df = 3 (P = 0.23); I² = 30%						
Test for overall effect: Z = 2.05 (P = 0.04)						
Test for subgroup differences: Chi² = 0.78, df = 1 (P = 0.38), I² = 0%						



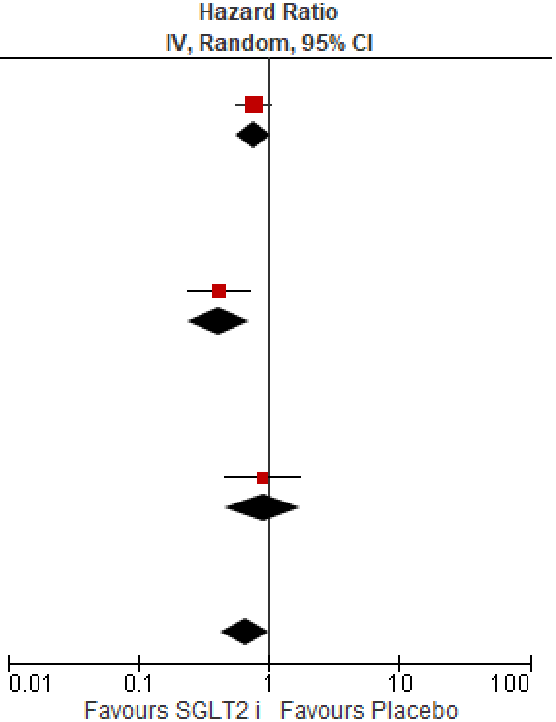
Hospitalisation for heart failure

Study or Subgroup	log[Hazard Ratio]	SE	SGLT2i Total	Placebo Total	Weight	Hazard Ratio IV, Fixed, 95% CI
2.7.1 eGFR (30-45)						
CANVAS (eGFR 30-45) - canagliflozin	-0.799	0.342	277	277	8.6%	0.45 [0.23, 0.88]
CREDENCE (eGFR 30-45)- canagliflozin	-0.35667	0.212959	657	656	22.1%	0.70 [0.46, 1.06]
EMPA-REG (eGFR 30-45) empagliflozin	-0.598	0.328	381	189	9.3%	0.55 [0.29, 1.05]
Subtotal (95% CI)			**1315**	**1122**	**39.9%**	**0.60 [0.44, 0.82]**
Heterogeneity: Chi² = 1.30, df = 2 (P = 0.52); I² = 0%						
Test for overall effect: Z = 3.21 (P = 0.001)						
2.7.2 eGFR (45-60)						
CANVAS (eGFR 45-60) - canagliflozin	-0.478	0.261	743	743	14.7%	0.62 [0.37, 1.03]
CREDENCE (eGFR 45-60)- canagliflozin	-0.84397	0.259839	640	639	14.8%	0.43 [0.26, 0.72]
EMPA-REG (eGFR 45-60) empagliflozin	-0.494	0.266	831	418	14.2%	0.61 [0.36, 1.03]
Subtotal (95% CI)			**2214**	**1800**	**43.7%**	**0.54 [0.40, 0.73]**
Heterogeneity: Chi² = 1.26, df = 2 (P = 0.53); I² = 0%						
Test for overall effect: Z = 4.01 (P < 0.0001)						
2.7.3 eGFR (60-90)						
CREDENCE (eGFR 60-90)- canagliflozin	-0.3285	0.247296	905	904	16.4%	0.72 [0.44, 1.17]
Subtotal (95% CI)			**905**	**904**	**16.4%**	**0.72 [0.44, 1.17]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.33 (P = 0.18)						
Total (95% CI)			**4434**	**3826**	**100.0%**	**0.59 [0.49, 0.72]**
Heterogeneity: Chi² = 3.50, df = 6 (P = 0.74); I² = 0%						
Test for overall effect: Z = 5.22 (P < 0.00001)						
Test for subgroup differences: Chi² = 0.94, df = 2 (P = 0.63), I² = 0%						



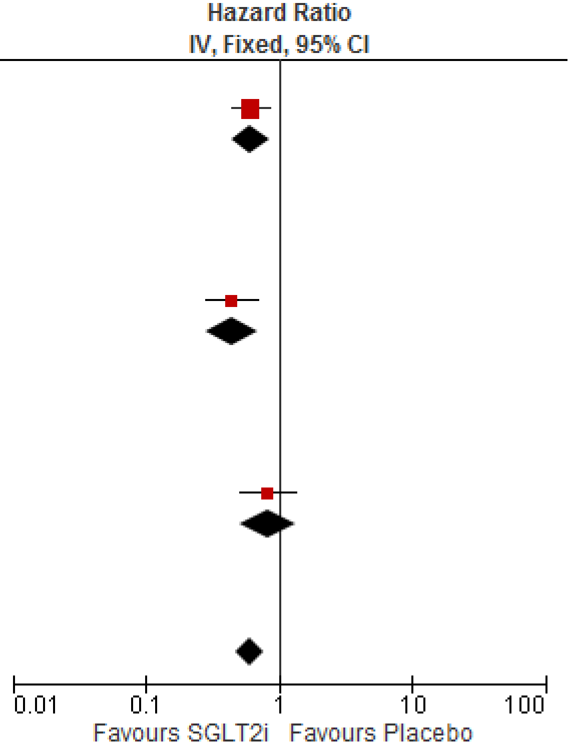
End stage kidney disease

Study or Subgroup	log[Hazard Ratio]	SE	SGLT-2 i Total	Placebo Total	Weight	Hazard Ratio IV, Random, 95% CI
2.8.1 eGFR (30-45)						
CREDENCE (eGFR 30-45)- canagliflozin	-0.27444	0.150451	657	656	46.0%	0.76 [0.57, 1.02]
Subtotal (95% CI)			**657**	**656**	**46.0%**	**0.76 [0.57, 1.02]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.82 (P = 0.07)						
2.8.2 eGFR 45-60						
CREDENCE (eGFR 45-60)- canagliflozin	-0.8916	0.27669	640	639	29.9%	0.41 [0.24, 0.71]
Subtotal (95% CI)			**640**	**639**	**29.9%**	**0.41 [0.24, 0.71]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 3.22 (P = 0.001)						
2.8.3 eGFR 45-60						
CREDENCE (eGFR 60-90)- canagliflozin	-0.11653	0.336442	905	904	24.1%	0.89 [0.46, 1.72]
Subtotal (95% CI)			**905**	**904**	**24.1%**	**0.89 [0.46, 1.72]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.35 (P = 0.73)						
Total (95% CI)			**2202**	**2199**	**100.0%**	**0.66 [0.43, 1.00]**
Heterogeneity: Tau² = 0.08; Chi² = 4.52, df = 2 (P = 0.10); I² = 56%						
Test for overall effect: Z = 1.96 (P = 0.05)						
Test for subgroup differences: Chi² = 4.52, df = 2 (P = 0.10), I² = 55.8%						



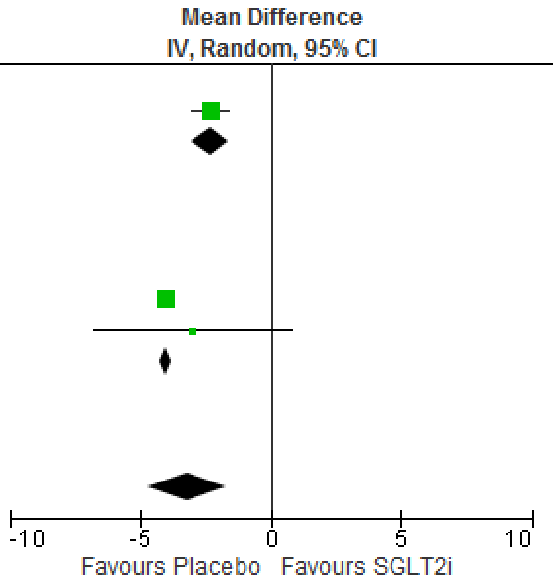
Doubling serum creatinine

Study or Subgroup	log[Hazard Ratio]	SE	SGLT2i Total	Placebo Total	Weight	Hazard Ratio IV, Fixed, 95% CI
2.9.1 eGFR (30-45)						
CREDENCE (eGFR 30-45)- canagliflozin	-0.4943	0.167975	657	656	49.6%	0.61 [0.44, 0.85]
Subtotal (95% CI)			**657**	**656**	**49.6%**	**0.61 [0.44, 0.85]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.94 (P = 0.003)						
2.9.2 eGFR (45-60)						
CREDENCE (eGFR 45-60)- canagliflozin	-0.82098	0.230077	640	639	26.4%	0.44 [0.28, 0.69]
Subtotal (95% CI)			**640**	**639**	**26.4%**	**0.44 [0.28, 0.69]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 3.57 (P = 0.0004)						
2.9.3 eGFR (60-90)						
CREDENCE (eGFR 60-90)- canagliflozin	-0.18633	0.241479	905	904	24.0%	0.83 [0.52, 1.33]
Subtotal (95% CI)			**905**	**904**	**24.0%**	**0.83 [0.52, 1.33]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.77 (P = 0.44)						
Total (95% CI)			**2202**	**2199**	**100.0%**	**0.60 [0.48, 0.76]**
Heterogeneity: Chi² = 3.63, df = 2 (P = 0.16); I² = 45%						
Test for overall effect: Z = 4.28 (P < 0.0001)						
Test for subgroup differences: Chi² = 3.63, df = 2 (P = 0.16), I² = 44.9%						



eGFR 6 months

Study or Subgroup	SGLT2i Mean	SGLT2i SD	SGLT2i Total	Placebo Mean	Placebo SD	Placebo Total	Weight	Mean Difference IV, Random, 95% CI
2.12.1 eGFR(30-45)								
CANVAS (eGFR 30-45) - canagliflozin (1)	37.8	3.0652	228	40.1	4.807	249	42.2%	-2.30 [-3.02, -1.58]
Subtotal (95% CI)			**228**			**249**	**42.2%**	**-2.30 [-3.02, -1.58]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 6.28 (P < 0.00001)								
2.12.2 eGFR (45-60)								
CANVAS (eGFR 45-60) - canagliflozin (2)	51	1.3913	746	55	2.5379	621	46.1%	-4.00 [-4.22, -3.78]
DERIVE (eGFR 45 to 60) - dapagliflozin (3)	-3.9	18.0831	142	-0.9	14.3625	140	11.7%	-3.00 [-6.81, 0.81]
Subtotal (95% CI)			**888**			**761**	**57.8%**	**-4.00 [-4.22, -3.77]**
Heterogeneity: Tau² = 0.00; Chi² = 0.26, df = 1 (P = 0.61); I² = 0%								
Test for overall effect: Z = 35.16 (P < 0.00001)								
Total (95% CI)			**1116**			**1010**	**100.0%**	**-3.17 [-4.67, -1.66]**
Heterogeneity: Tau² = 1.27; Chi² = 19.85, df = 2 (P < 0.0001); I² = 90%								
Test for overall effect: Z = 4.12 (P < 0.0001)								
Test for subgroup differences: Chi² = 19.59, df = 1 (P < 0.00001), I² = 94.9%								

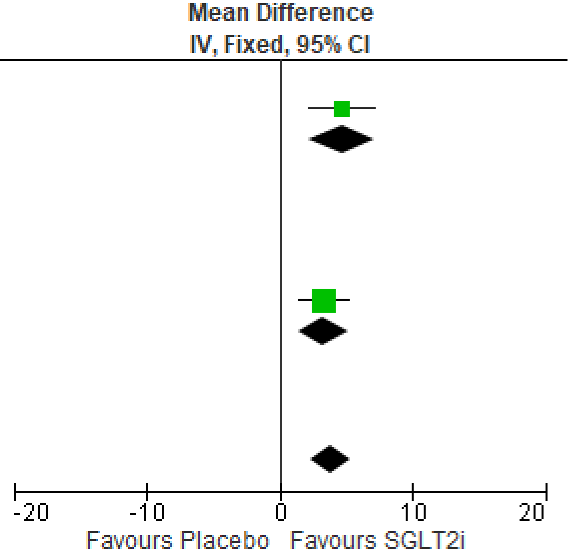


Footnotes
(1) Read from graph
(2) Read from graph
(3) Read from graph. 24 weeks follow up.

4) eGFR 6 months (ml/min/1.73m2, higher values are better)

eGFR last available data point > 2years

Study or Subgroup	SGLT2i Mean	SGLT2i SD	SGLT2i Total	Placebo Mean	Placebo SD	Placebo Total	Weight	Mean Difference IV, Fixed, 95% CI
2.13.1 eGFR (30-45)								
CANVAS (eGFR 30-45) - canagliflozin (1)	40.5	4.9928	45	35.8	4.2789	17	36.5%	4.70 [2.20, 7.20]
Subtotal (95% CI)			**45**			**17**	**36.5%**	**4.70 [2.20, 7.20]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 3.68 (P = 0.0002)								
2.13.2 eGFR (45-60)								
CANVAS (eGFR 45-60) - canagliflozin (2)	53.8	7.8305	166	50.5	6.9963	86	63.5%	3.30 [1.40, 5.20]
Subtotal (95% CI)			**166**			**86**	**63.5%**	**3.30 [1.40, 5.20]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 3.41 (P = 0.0007)								
Total (95% CI)			**211**			**103**	**100.0%**	**3.81 [2.30, 5.32]**
Heterogeneity: Chi² = 0.76, df = 1 (P = 0.38); I² = 0%								
Test for overall effect: Z = 4.94 (P < 0.00001)								
Test for subgroup differences: Chi² = 0.76, df = 1 (P = 0.38), I² = 0%								

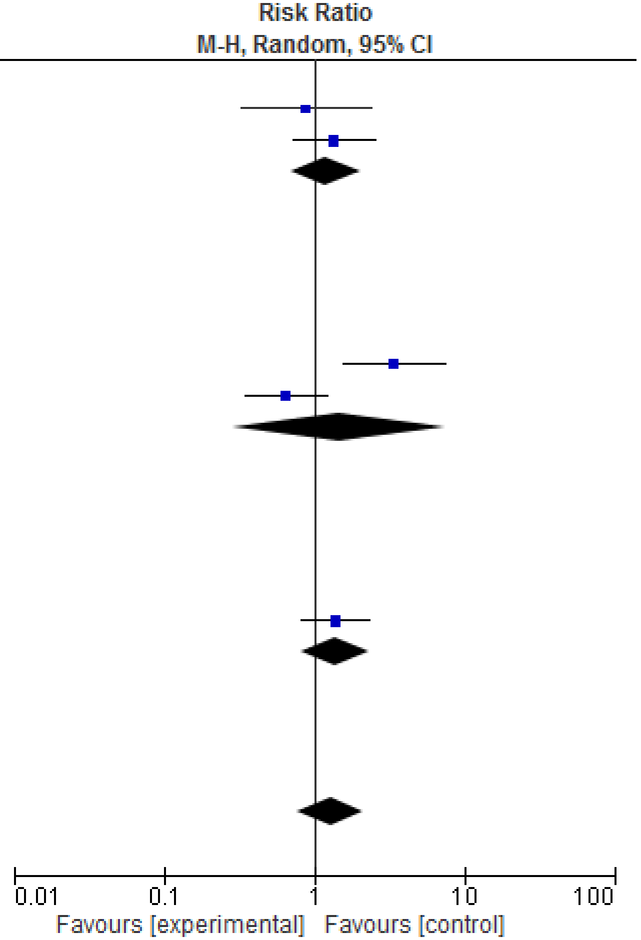


Footnotes
(1) Read from graph
(2) Read from graph

3)eGFR >2years (ml/min/1.73m2, higher values are better)

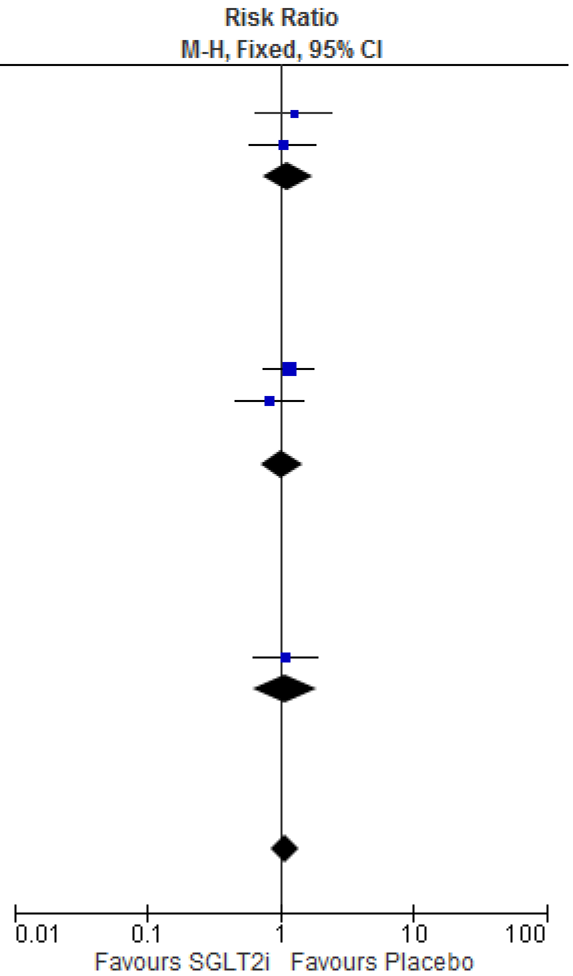
Amputation

Study or Subgroup	SGLT2i Events	SGLT2i Total	Placebo Events	Placebo Total	Weight	Risk Ratio M-H, Random, 95% CI
2.17.1 eGFR (30-45)						
CANVAS (eGFR 30-45) - canagliflozin	7	277	8	277	14.4%	0.88 [0.32, 2.38]
CREDENCE (eGFR 30-45)- canagliflozin	23	655	17	656	21.9%	1.36 [0.73, 2.51]
Subtotal (95% CI)		**932**		**933**	**36.3%**	**1.20 [0.71, 2.03]**
Total events	30		25			
Heterogeneity: Tau² = 0.00; Chi² = 0.53, df = 1 (P = 0.47); I² = 0%						
Test for overall effect: Z = 0.68 (P = 0.49)						
2.17.2 eGFR (45-60)						
CANVAS (eGFR 45-60) - canagliflozin	27	743	8	743	18.3%	3.38 [1.54, 7.38]
CREDENCE (eGFR 45-60)- canagliflozin	15	640	23	638	21.4%	0.65 [0.34, 1.23]
Subtotal (95% CI)		**1383**		**1381**	**39.7%**	**1.46 [0.29, 7.37]**
Total events	42		31			
Heterogeneity: Tau² = 1.23; Chi² = 10.26, df = 1 (P = 0.001); I² = 90%						
Test for overall effect: Z = 0.46 (P = 0.65)						
2.17.3 eGFR (60-90)						
CREDENCE (eGFR 60-90)- canagliflozin	32	905	23	903	24.0%	1.39 [0.82, 2.35]
Subtotal (95% CI)		**905**		**903**	**24.0%**	**1.39 [0.82, 2.35]**
Total events	32		23			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.22 (P = 0.22)						
Total (95% CI)		**3220**		**3217**	**100.0%**	**1.29 [0.78, 2.15]**
Total events	104		79			
Heterogeneity: Tau² = 0.21; Chi² = 10.90, df = 4 (P = 0.03); I² = 63%						
Test for overall effect: Z = 0.99 (P = 0.32)						
Test for subgroup differences: Chi² = 0.17, df = 2 (P = 0.92), I² = 0%						



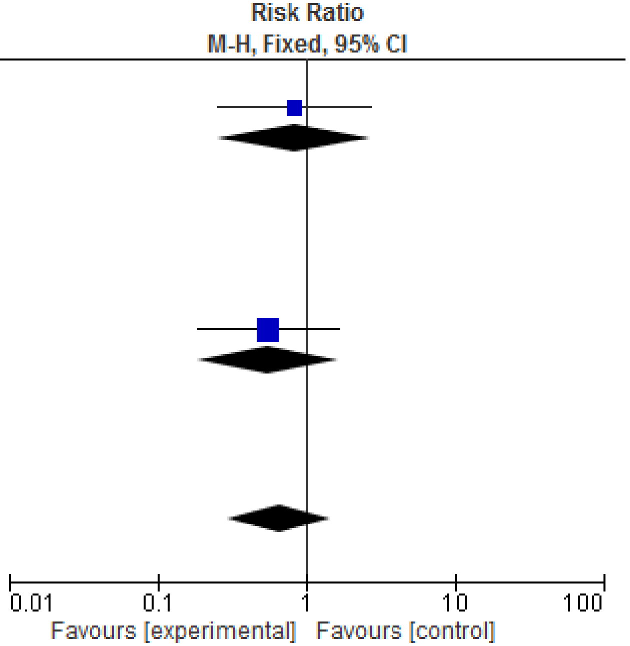
Fracture

Study or Subgroup	SGLT2i Events	SGLT2i Total	Placebo Events	Placebo Total	Weight	Risk Ratio M-H, Fixed, 95% CI
2.18.1 eGFR (30-45)						
CANVAS (eGFR 30-45) - canagliflozin	19	277	15	277	12.6%	1.27 [0.66, 2.44]
CREDENCE (eGFR 30-45)- canagliflozin	23	655	22	656	18.5%	1.05 [0.59, 1.86]
Subtotal (95% CI)		**932**		**933**	**31.1%**	**1.14 [0.74, 1.75]**
Total events	42		37			
Heterogeneity: Chi² = 0.18, df = 1 (P = 0.67); I² = 0%						
Test for overall effect: Z = 0.58 (P = 0.56)						
2.18.2 eGFR (45-60)						
CANVAS (eGFR 45-60) - canagliflozin	42	743	36	743	30.2%	1.17 [0.76, 1.80]
CREDENCE (eGFR 45-60)- canagliflozin	19	640	23	638	19.4%	0.82 [0.45, 1.50]
DERIVE (eGFR 45 to 60) - dapagliflozin	0	160	0	161		Not estimable
Subtotal (95% CI)		**1543**		**1542**	**49.6%**	**1.03 [0.73, 1.47]**
Total events	61		59			
Heterogeneity: Chi² = 0.86, df = 1 (P = 0.36); I² = 0%						
Test for overall effect: Z = 0.18 (P = 0.86)						
2.18.3 eGFR (60-90)						
CREDENCE (eGFR 60-90)- canagliflozin	25	905	23	903	19.3%	1.08 [0.62, 1.90]
Subtotal (95% CI)		**905**		**903**	**19.3%**	**1.08 [0.62, 1.90]**
Total events	25		23			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.28 (P = 0.78)						
Total (95% CI)		**3380**		**3378**	**100.0%**	**1.07 [0.84, 1.37]**
Total events	128		119			
Heterogeneity: Chi² = 1.15, df = 4 (P = 0.89); I² = 0%						
Test for overall effect: Z = 0.58 (P = 0.56)						
Test for subgroup differences: Chi² = 0.11, df = 2 (P = 0.94), I² = 0%						



Acute Kidney Injury

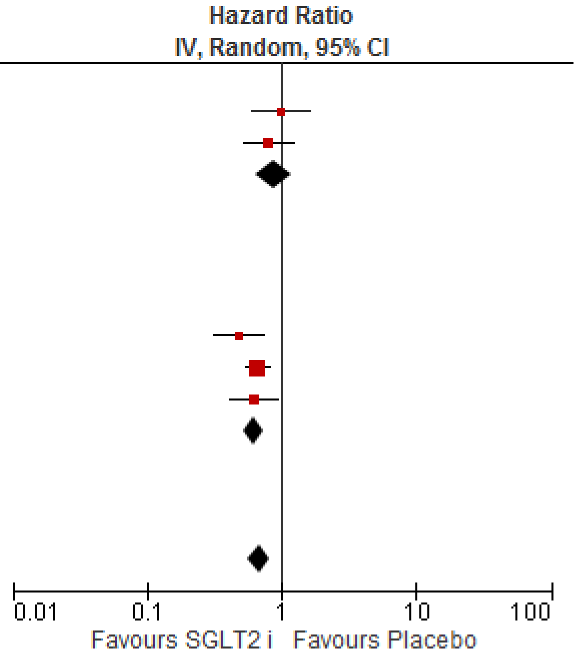
Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk Ratio M-H, Fixed, 95% CI
2.19.1 eGFR (30-45)						
CANVAS (eGFR 30-45) - canagliflozin	5	277	6	277	40.0%	0.83 [0.26, 2.70]
Subtotal (95% CI)		**277**		**277**	**40.0%**	**0.83 [0.26, 2.70]**
Total events	5		6			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.30 (P = 0.76)						
2.19.2 eGFR (45-60)						
CANVAS (eGFR 45-60) - canagliflozin	5	743	9	743	60.0%	0.56 [0.19, 1.65]
Subtotal (95% CI)		**743**		**743**	**60.0%**	**0.56 [0.19, 1.65]**
Total events	5		9			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.06 (P = 0.29)						
Total (95% CI)		**1020**		**1020**	**100.0%**	**0.67 [0.30, 1.48]**
Total events	10		15			
Heterogeneity: Chi² = 0.25, df = 1 (P = 0.62); I² = 0%						
Test for overall effect: Z = 1.00 (P = 0.32)						
Test for subgroup differences: Chi² = 0.25, df = 1 (P = 0.62), I² = 0%						



SGLT2 vs Placebo ACR stratification

Renal composite – End stage kidney disease, doubling serum creatinine, renal death

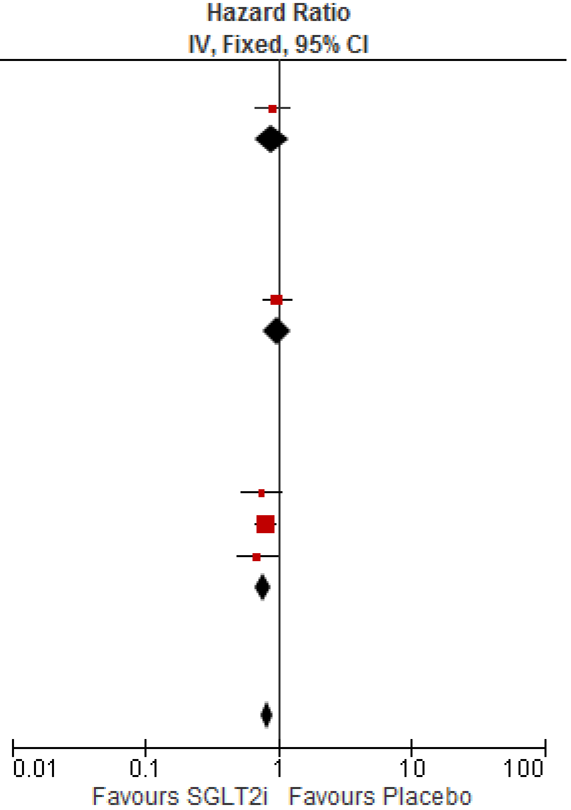
Study or Subgroup	log[Hazard Ratio]	SE	SGLT-2 i Total	Placebo Total	Weight	Hazard Ratio IV, Random, 95% CI
3.1.2 A2						
CANVAS A2	-0.0202	0.250212	1322	944	12.6%	0.98 [0.60, 1.60]
VERTIS CV (A2) - ertugliflozin (1)	-0.223	0.21	5499	2747	16.7%	0.80 [0.53, 1.21]
Subtotal (95% CI)			**6821**	**3691**	**29.4%**	**0.87 [0.63, 1.19]**
Heterogeneity: Tau² = 0.00; Chi² = 0.39, df = 1 (P = 0.53); I² = 0%						
Test for overall effect: Z = 0.87 (P = 0.39)						
3.1.3 A3						
CANVAS A3	-0.73397	0.221959	406	354	15.4%	0.48 [0.31, 0.74]
CREDENCE (A3) - canagliflozin	-0.416	0.108	2202	2199	39.0%	0.66 [0.53, 0.82]
VERTIS CV (A3) - ertugliflozin	-0.478	0.214	513	242	16.3%	0.62 [0.41, 0.94]
Subtotal (95% CI)			**3121**	**2795**	**70.6%**	**0.62 [0.52, 0.74]**
Heterogeneity: Tau² = 0.00; Chi² = 1.66, df = 2 (P = 0.44); I² = 0%						
Test for overall effect: Z = 5.39 (P < 0.00001)						
Total (95% CI)			**9942**	**6486**	**100.0%**	**0.68 [0.56, 0.82]**
Heterogeneity: Tau² = 0.01; Chi² = 5.43, df = 4 (P = 0.25); I² = 26%						
Test for overall effect: Z = 4.02 (P < 0.0001)						
Test for subgroup differences: Chi² = 3.39, df = 1 (P = 0.07), I² = 70.5%						



Footnotes
(1) Totals given are for whole trial population. For A2 n=2492 across placebo/intervention and A3 n=755

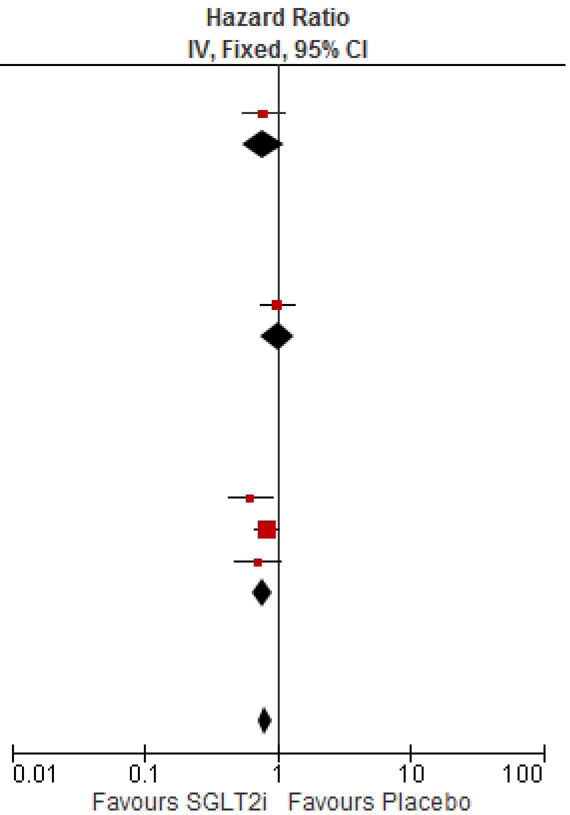
Cardiovascular composite: 3-point MACE

Study or Subgroup	log[Hazard Ratio]	SE	SGLT2i Total	Placebo Total	Weight	Hazard Ratio IV, Fixed, 95% CI
3.2.1 A1&A2						
EMPA-REG (A1&A2) empagliflozin	-0.105	0.15289	850	440	14.4%	0.90 [0.67, 1.21]
Subtotal (95% CI)			**850**	**440**	**14.4%**	**0.90 [0.67, 1.21]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.69 (P = 0.49)						
3.2.2 A2						
CANVAS A2	-0.0202	0.126934	1322	944	20.9%	0.98 [0.76, 1.26]
Subtotal (95% CI)			**1322**	**944**	**20.9%**	**0.98 [0.76, 1.26]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.16 (P = 0.87)						
3.2.3 A3						
CANVAS A3	-0.28768	0.176823	406	354	10.8%	0.75 [0.53, 1.06]
CREDENCE (A3) - canagliflozin	-0.223	0.089	2202	2199	42.5%	0.80 [0.67, 0.95]
EMPA-REG (A3) empaglaflozin	-0.371	0.172	1498	752	11.4%	0.69 [0.49, 0.97]
Subtotal (95% CI)			**4106**	**3305**	**64.7%**	**0.77 [0.67, 0.89]**
Heterogeneity: Chi^2 = 0.61, df = 2 (P = 0.74); I^2 = 0%						
Test for overall effect: Z = 3.60 (P = 0.0003)						
Total (95% CI)			**6278**	**4689**	**100.0%**	**0.83 [0.74, 0.93]**
Heterogeneity: Chi^2 = 3.65, df = 4 (P = 0.46); I^2 = 0%						
Test for overall effect: Z = 3.23 (P = 0.001)						
Test for subgroup differences: Chi^2 = 3.03, df = 2 (P = 0.22), I^2 = 34.1%						



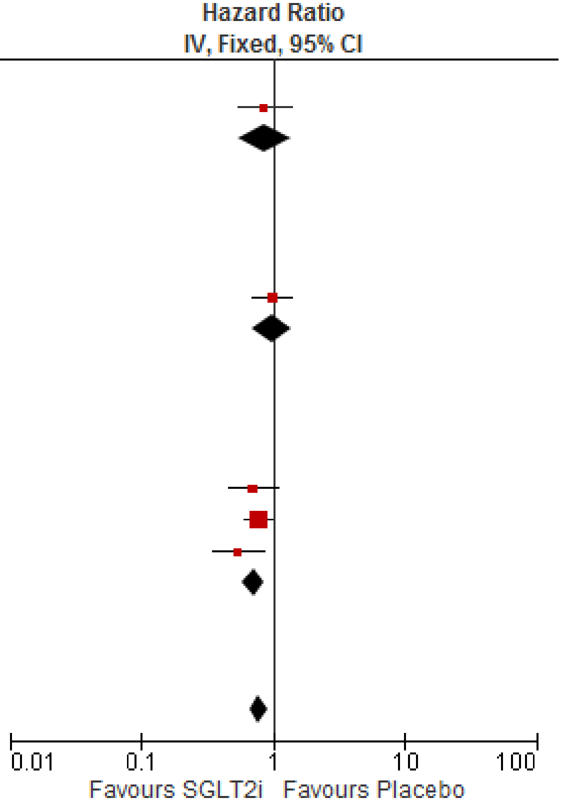
All-cause mortality

Study or Subgroup	log[Hazard Ratio]	SE	SGLT2i Total	Placebo Total	Weight	Hazard Ratio IV, Fixed, 95% CI
3.3.1 A1&A2						
EMPA-REG (A1&A2) empagliflozin	-0.248	0.181	850	440	13.9%	0.78 [0.55, 1.11]
Subtotal (95% CI)			**850**	**440**	**13.9%**	**0.78 [0.55, 1.11]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.37 (P = 0.17)						
3.3.2 A2						
CANVAS A2	0	0.151473	1322	944	19.8%	1.00 [0.74, 1.35]
Subtotal (95% CI)			**1322**	**944**	**19.8%**	**1.00 [0.74, 1.35]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.00 (P = 1.00)						
3.3.3 A3						
CANVAS A3	-0.46204	0.194028	406	354	12.1%	0.63 [0.43, 0.92]
CREDENCE (A3) - canagliflozin	-0.186	0.103	2202	2199	42.8%	0.83 [0.68, 1.02]
EMPA-REG (A3) empaglaflozin	-0.34249	0.199684	509	260	11.4%	0.71 [0.48, 1.05]
Subtotal (95% CI)			**3117**	**2813**	**66.3%**	**0.77 [0.65, 0.90]**
Heterogeneity: Chi² = 1.77, df = 2 (P = 0.41); I² = 0%						
Test for overall effect: Z = 3.18 (P = 0.001)						
Total (95% CI)			**5289**	**4197**	**100.0%**	**0.81 [0.71, 0.93]**
Heterogeneity: Chi² = 4.15, df = 4 (P = 0.39); I² = 4%						
Test for overall effect: Z = 3.10 (P = 0.002)						
Test for subgroup differences: Chi² = 2.38, df = 2 (P = 0.30), I² = 15.9%						



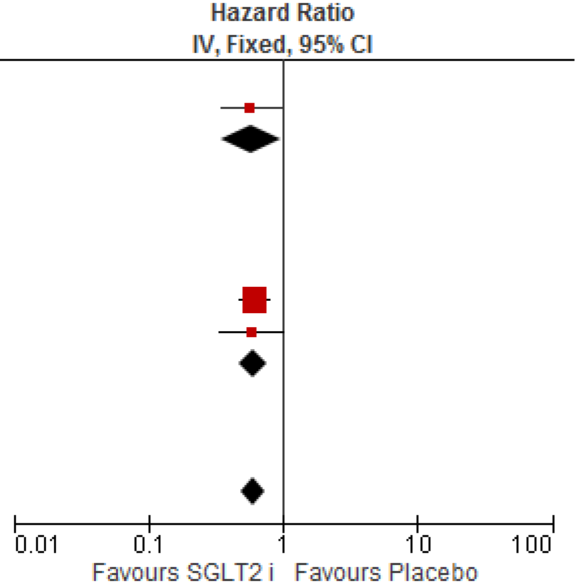
Cardiovascular death

Study or Subgroup	log[Hazard Ratio]	SE	SGLT2 i Total	Placebo Total	Weight	Hazard Ratio IV, Fixed, 95% CI
3.4.1 A1&A2						
EMPA-REG (A1&A2) empagliflozin	-0.151	0.237	850	440	11.8%	0.86 [0.54, 1.37]
Subtotal (95% CI)			**850**	**440**	**11.8%**	**0.86 [0.54, 1.37]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.64 (P = 0.52)						
3.4.2 A2						
CANVAS A2	-0.0202	0.18231	1322	944	19.9%	0.98 [0.69, 1.40]
Subtotal (95% CI)			**1322**	**944**	**19.9%**	**0.98 [0.69, 1.40]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.11 (P = 0.91)						
3.4.3 A3						
CANVAS A3	-0.35667494	0.220961	406	354	13.6%	0.70 [0.45, 1.08]
CREDENCE (A3) - canagliflozin	-0.248	0.126	2202	2199	41.7%	0.78 [0.61, 1.00]
EMPA-REG (A3) empaglaflozin	-0.616	0.226	509	260	13.0%	0.54 [0.35, 0.84]
Subtotal (95% CI)			**3117**	**2813**	**68.3%**	**0.71 [0.59, 0.86]**
Heterogeneity: Chi² = 2.03, df = 2 (P = 0.36); I² = 1%						
Test for overall effect: Z = 3.45 (P = 0.0006)						
Total (95% CI)			**5289**	**4197**	**100.0%**	**0.78 [0.66, 0.91]**
Heterogeneity: Chi² = 4.62, df = 4 (P = 0.33); I² = 13%						
Test for overall effect: Z = 3.12 (P = 0.002)						
Test for subgroup differences: Chi² = 2.59, df = 2 (P = 0.27), I² = 22.7%						



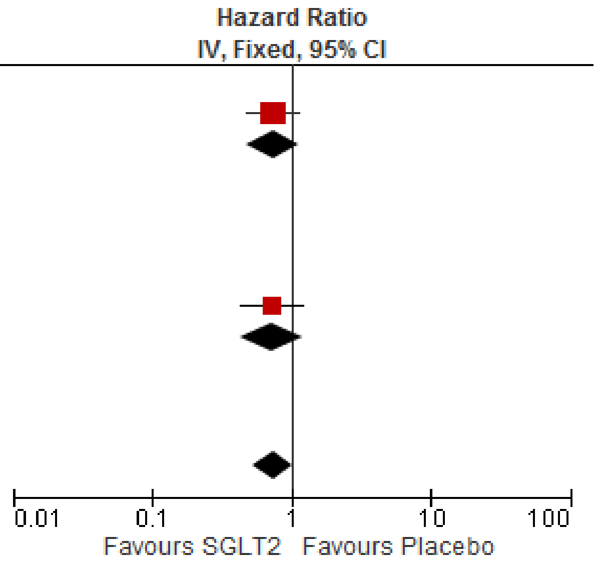
Hospitalisation for heart failure

Study or Subgroup	log[Hazard Ratio]	SE	SGLT2i Total	Placebo Total	Weight	Hazard Ratio IV, Fixed, 95% CI
3.7.1 A1&A2						
EMPA-REG (A1&A2) empagliflozin	-0.562	0.262	850	440	17.8%	0.57 [0.34, 0.95]
Subtotal (95% CI)			**850**	**440**	**17.8%**	**0.57 [0.34, 0.95]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.15 (P = 0.03)						
3.7.2 A3						
CREDENCE (A3) - canagliflozin	-0.494	0.136	2202	2199	65.9%	0.61 [0.47, 0.80]
EMPA-REG (A3) empaglaflozin	-0.545	0.273	509	260	16.4%	0.58 [0.34, 0.99]
Subtotal (95% CI)			**2711**	**2459**	**82.2%**	**0.60 [0.48, 0.77]**
Heterogeneity: Chi² = 0.03, df = 1 (P = 0.87); I² = 0%						
Test for overall effect: Z = 4.14 (P < 0.0001)						
Total (95% CI)			**3561**	**2899**	**100.0%**	**0.60 [0.48, 0.74]**
Heterogeneity: Chi² = 0.07, df = 2 (P = 0.97); I² = 0%						
Test for overall effect: Z = 4.66 (P < 0.00001)						
Test for subgroup differences: Chi² = 0.04, df = 1 (P = 0.84), I² = 0%						



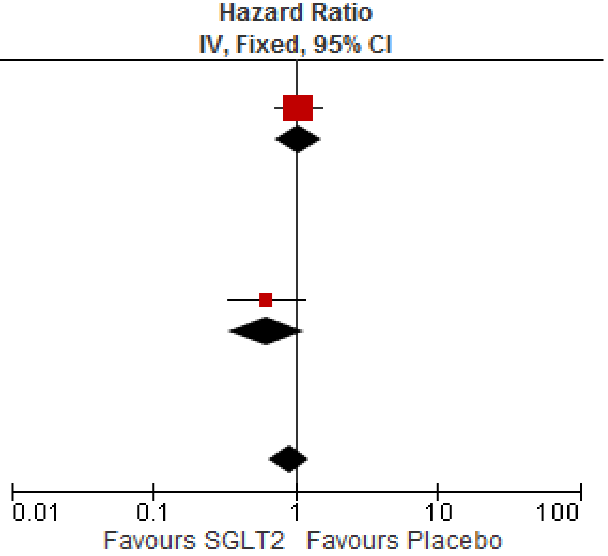
Hospitalised for or fatal heart failure

Study or Subgroup	log[Hazard Ratio]	SE	SGLT2 Total	PLacebo Total	Weight	Hazard Ratio IV, Fixed, 95% CI
3.10.1 A2						
CANVAS A2	-0.30111	0.222891	1322	944	58.8%	0.74 [0.48, 1.15]
Subtotal (95% CI)			**1322**	**944**	**58.8%**	**0.74 [0.48, 1.15]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.35 (P = 0.18)						
3.10.2 A3						
CANVAS A3	-0.31471	0.266026	406	354	41.2%	0.73 [0.43, 1.23]
Subtotal (95% CI)			**406**	**354**	**41.2%**	**0.73 [0.43, 1.23]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.18 (P = 0.24)						
Total (95% CI)			**1728**	**1298**	**100.0%**	**0.74 [0.53, 1.03]**
Heterogeneity: Chi² = 0.00, df = 1 (P = 0.97); I² = 0%						
Test for overall effect: Z = 1.80 (P = 0.07)						
Test for subgroup differences: Chi² = 0.00, df = 1 (P = 0.97), I² = 0%						



Fatal/non-fatal MI

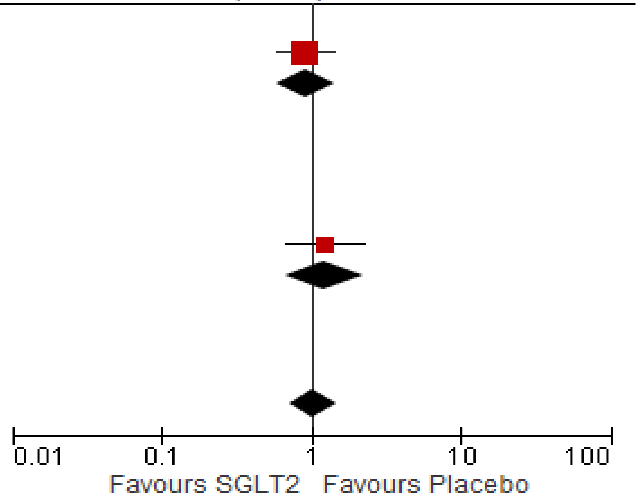
Study or Subgroup	log[Hazard Ratio]	SE	SGLT2 Total	Placebo Total	Weight	Hazard Ratio IV, Fixed, 95% CI
3.5.1 A2						
CANVAS A2	0.04879	0.20081	1322	944	71.8%	1.05 [0.71, 1.56]
Subtotal (95% CI)			**1322**	**944**	**71.8%**	**1.05 [0.71, 1.56]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.24 (P = 0.81)						
3.5.2 A3						
CANVAS A3	-0.47804	0.320684	406	354	28.2%	0.62 [0.33, 1.16]
Subtotal (95% CI)			**406**	**354**	**28.2%**	**0.62 [0.33, 1.16]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.49 (P = 0.14)						
Total (95% CI)			**1728**	**1298**	**100.0%**	**0.91 [0.65, 1.26]**
Heterogeneity: Chi² = 1.94, df = 1 (P = 0.16); I² = 48%						
Test for overall effect: Z = 0.59 (P = 0.56)						
Test for subgroup differences: Chi² = 1.94, df = 1 (P = 0.16), I² = 48.4%						



Fatal/non-fatal stroke

Study or Subgroup	log[Hazard Ratio]	SE	SGLT2 Total	Placebo Total	Weight	Hazard Ratio IV, Fixed, 95% CI
3.6.1 A2						
CANVAS A2	-0.08338	0.23114	1322	944	64.3%	0.92 [0.58, 1.45]
Subtotal (95% CI)			**1322**	**944**	**64.3%**	**0.92 [0.58, 1.45]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.36 (P = 0.72)						
3.6.2 A3						
CANVAS A3	0.207014	0.310164	406	354	35.7%	1.23 [0.67, 2.26]
Subtotal (95% CI)			**406**	**354**	**35.7%**	**1.23 [0.67, 2.26]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.67 (P = 0.50)						
Total (95% CI)			**1728**	**1298**	**100.0%**	**1.02 [0.71, 1.47]**
Heterogeneity: Chi² = 0.56, df = 1 (P = 0.45); I² = 0%						
Test for overall effect: Z = 0.11 (P = 0.91)						
Test for subgroup differences: Chi² = 0.56, df = 1 (P = 0.45), I² = 0%						

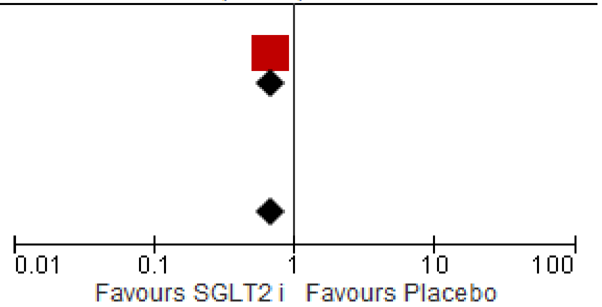
Hazard Ratio
IV, Fixed, 95% CI



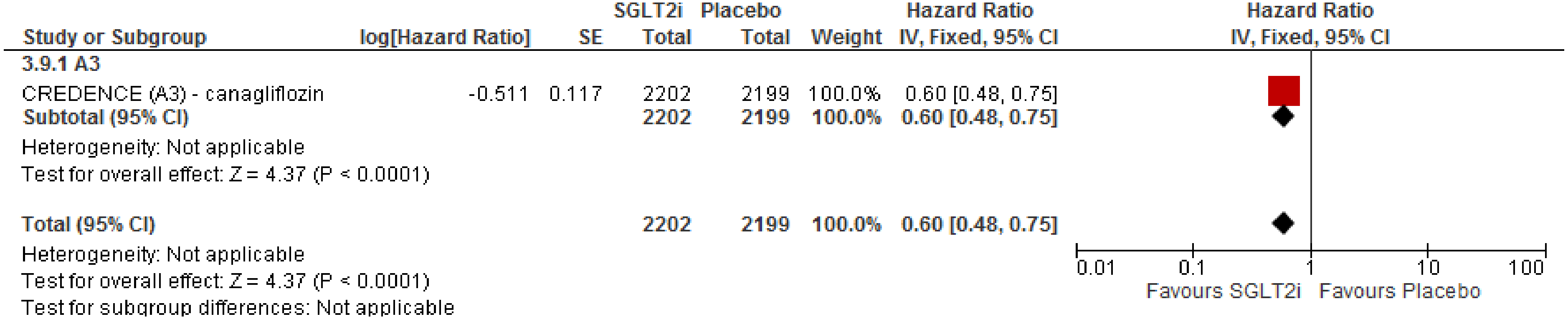
End stage kidney disease

Study or Subgroup	log[Hazard Ratio]	SE	SGLT-2 i Total	Placebo Total	Weight	Hazard Ratio IV, Fixed, 95% CI
3.8.1 A3						
CREDENCE (A3) - canagliflozin	-0.386	0.119	2202	2199	100.0%	0.68 [0.54, 0.86]
Subtotal (95% CI)			**2202**	**2199**	**100.0%**	**0.68 [0.54, 0.86]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 3.24 (P = 0.001)						
Total (95% CI)			**2202**	**2199**	**100.0%**	**0.68 [0.54, 0.86]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 3.24 (P = 0.001)						
Test for subgroup differences: Not applicable						

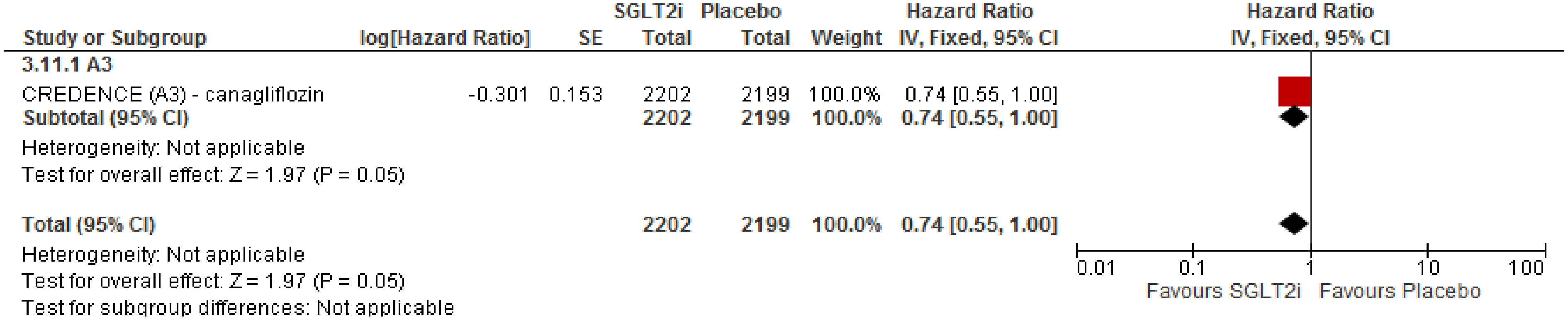
Hazard Ratio
IV, Fixed, 95% CI



Doubling of serum creatinine

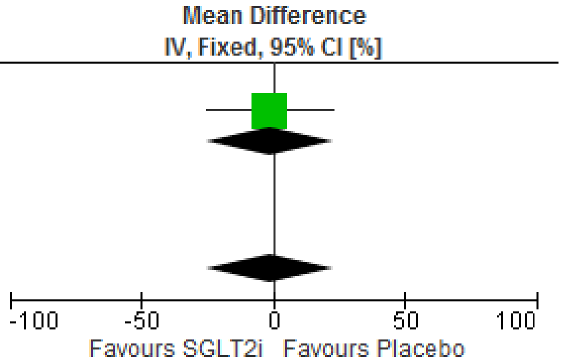


Dialysis



Percentage change from baseline UACR 6 months

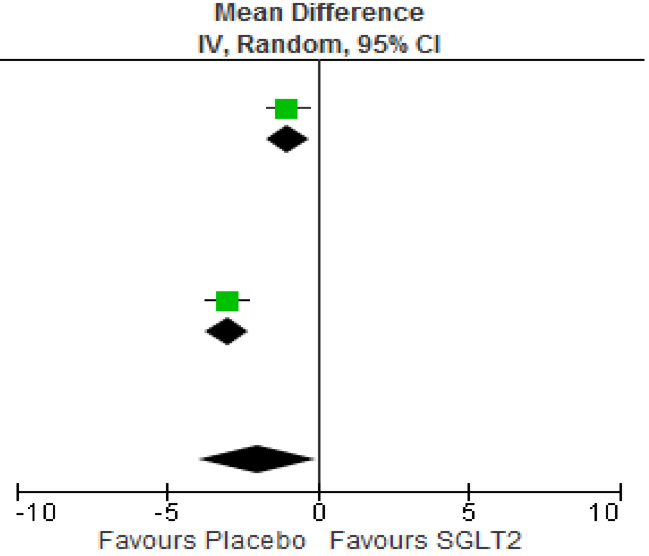
Study or Subgroup	SGLT2i Mean [%]	SGLT2i SD [%]	SGLT2i Total	Placebo Mean [%]	Placebo SD [%]	Placebo Total	Weight	Mean Difference IV, Fixed, 95% CI [%]
3.14.1 A3								
DELIGHT (A3) - dapagliflozin	-2	104.6506	148	-1	103.2027	144	100.0%	-1.00 [-24.84, 22.84]
Subtotal (95% CI)			**148**			**144**	**100.0%**	**-1.00 [-24.84, 22.84]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.08 (P = 0.93)								
Total (95% CI)			**148**			**144**	**100.0%**	**-1.00 [-24.84, 22.84]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.08 (P = 0.93)								
Test for subgroup differences: Not applicable								



1) Percentage change from baseline uACR (%, lower values are better)

eGFR 6 months

Study or Subgroup	SGLT2 Mean	SGLT2 SD	SGLT2 Total	Placebo Mean	Placebo SD	Placebo Total	Weight	Mean Difference IV, Random, 95% CI
3.12.1 A2								
CANVAS A2	72	8.9454	1232	73	7.54	876	50.0%	-1.00 [-1.71, -0.29]
Subtotal (95% CI)			**1232**			**876**	**50.0%**	**-1.00 [-1.71, -0.29]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.78 (P = 0.006)								
3.12.2 A3								
CANVAS A3	61	4.9109	373	64	4.5818	325	50.0%	-3.00 [-3.70, -2.30]
Subtotal (95% CI)			**373**			**325**	**50.0%**	**-3.00 [-3.70, -2.30]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 8.34 (P < 0.00001)								
Total (95% CI)			**1605**			**1201**	**100.0%**	**-2.00 [-3.96, -0.04]**
Heterogeneity: Tau² = 1.87; Chi² = 15.44, df = 1 (P < 0.0001); I² = 94%								
Test for overall effect: Z = 2.00 (P = 0.05)								
Test for subgroup differences: Chi² = 15.44, df = 1 (P < 0.0001), I² = 93.5%								

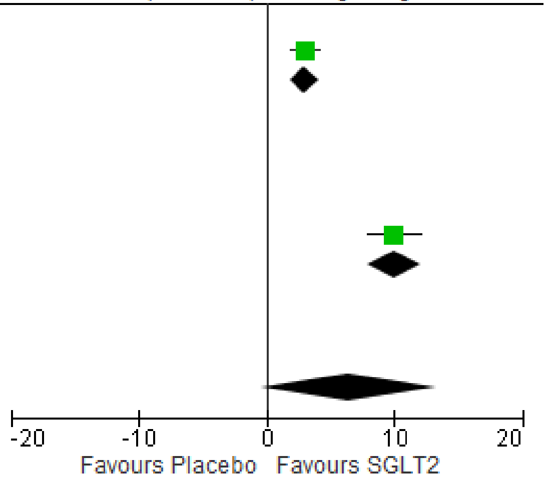


1 Numbers read from graph. eGFR 6 months (ml/min/1.73m2, higher values are better)

eGFR last available data point >2years

Study or Subgroup	SGLT2 Mean [eGFR]	SGLT2 SD [eGFR]	SGLT2 Total	Placebo Mean [eGFR]	Placebo SD [eGFR]	Placebo Total	Weight	Mean Difference IV, Random, 95% CI [eGFR]
3.13.1 A2								
CANVAS A2	71	4.7492	349	68	6.3843	159	50.9%	3.00 [1.89, 4.11]
Subtotal (95% CI)			**349**			**159**	**50.9%**	**3.00 [1.89, 4.11]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 5.30 (P < 0.00001)								
3.13.2 A3								
CANVAS A3	52	3.9707	63	42	4.9516	26	49.1%	10.00 [7.86, 12.14]
Subtotal (95% CI)			**63**			**26**	**49.1%**	**10.00 [7.86, 12.14]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 9.15 (P < 0.00001)								
Total (95% CI)			**412**			**185**	**100.0%**	**6.44 [-0.42, 13.30]**
Heterogeneity: Tau² = 23.74; Chi² = 32.36, df = 1 (P < 0.00001); I² = 97%								
Test for overall effect: Z = 1.84 (P = 0.07)								
Test for subgroup differences: Chi² = 32.36, df = 1 (P < 0.00001), I² = 96.9%								

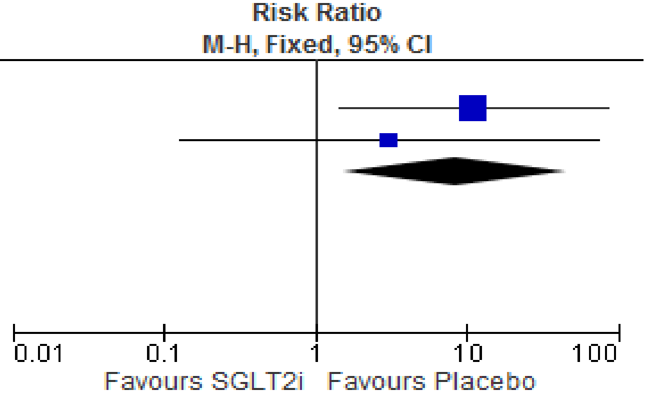
Mean Difference
IV, Random, 95% CI [eGFR]



[1] Numbers read from graph. eGFR > 2 years (ml/min/1.73m2, higher values are better)

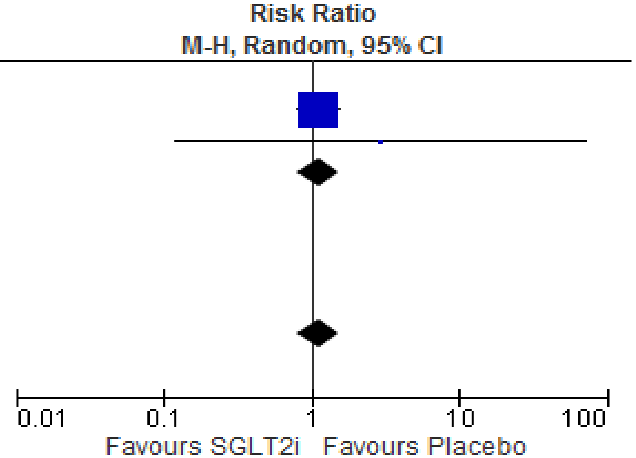
Diabetic ketoacidosis

Study or Subgroup	SGLT2i Events	SGLT2i Total	Placebo Events	Placebo Total	Weight	Risk Ratio M-H, Fixed, 95% CI
3.16.1 A3						
CREDENCE (A3) - canagliflozin	11	2200	1	2197	66.9%	10.98 [1.42, 85.01]
DELIGHT (A3) - dapagliflozin	1	145	0	148	33.1%	3.06 [0.13, 74.55]
Subtotal (95% CI)		**2345**		**2345**	**100.0%**	**8.36 [1.54, 45.52]**
Total events	12		1			
Heterogeneity: Chi² = 0.45, df = 1 (P = 0.50); I² = 0%						
Test for overall effect: Z = 2.46 (P = 0.01)						
Test for subgroup differences: Not applicable						



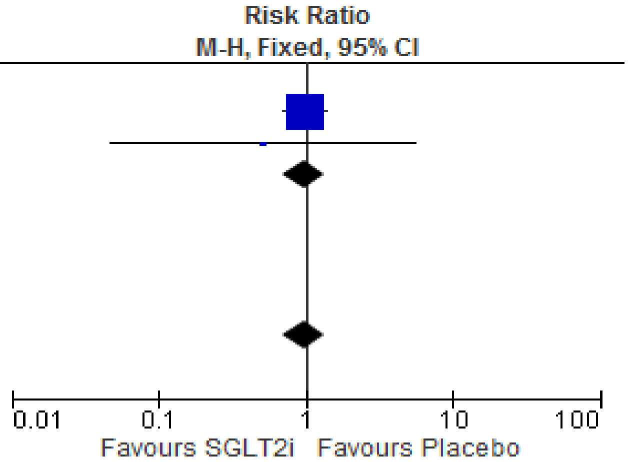
Amputation

Study or Subgroup	SGLT2i Events	SGLT2i Total	Placebo Events	Placebo Total	Weight	Risk Ratio M-H, Random, 95% CI
3.17.1 A3						
CREDENCE (A3) - canagliflozin	70	2200	63	2197	98.9%	1.11 [0.79, 1.55]
DELIGHT (A3) - dapagliflozin	1	148	0	145	1.1%	2.94 [0.12, 71.57]
Subtotal (95% CI)		**2348**		**2342**	**100.0%**	**1.12 [0.80, 1.57]**
Total events	71		63			
Heterogeneity: Tau² = 0.00; Chi² = 0.35, df = 1 (P = 0.55); I² = 0%						
Test for overall effect: Z = 0.67 (P = 0.50)						
Total (95% CI)		**2348**		**2342**	**100.0%**	**1.12 [0.80, 1.57]**
Total events	71		63			
Heterogeneity: Tau² = 0.00; Chi² = 0.35, df = 1 (P = 0.55); I² = 0%						
Test for overall effect: Z = 0.67 (P = 0.50)						
Test for subgroup differences: Not applicable						



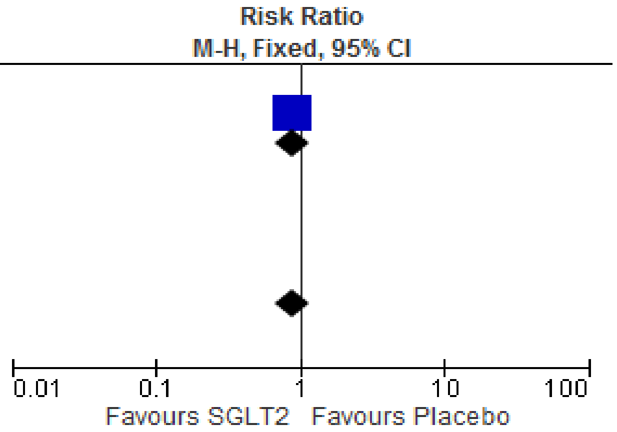
Fracture

Study or Subgroup	SGLT2i Events	SGLT2i Total	Placebo Events	Placebo Total	Weight	Risk Ratio M-H, Fixed, 95% CI
3.18.1 A3						
CREDENCE (A3) - canagliflozin	67	2200	68	2197	97.2%	0.98 [0.71, 1.37]
DELIGHT (A3) - dapagliflozin	1	145	2	148	2.8%	0.51 [0.05, 5.57]
Subtotal (95% CI)		**2345**		**2345**	**100.0%**	**0.97 [0.70, 1.35]**
Total events	68		70			
Heterogeneity: Chi² = 0.28, df = 1 (P = 0.59); I² = 0%						
Test for overall effect: Z = 0.18 (P = 0.86)						
Total (95% CI)		**2345**		**2345**	**100.0%**	**0.97 [0.70, 1.35]**
Total events	68		70			
Heterogeneity: Chi² = 0.28, df = 1 (P = 0.59); I² = 0%						
Test for overall effect: Z = 0.18 (P = 0.86)						
Test for subgroup differences: Not applicable						



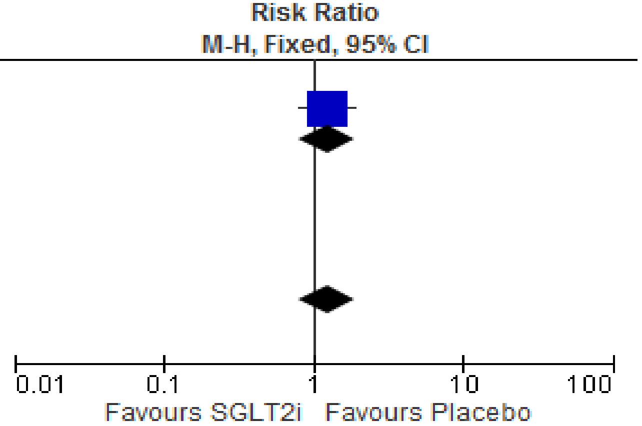
Acute Kidney Injury

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk Ratio M-H, Fixed, 95% CI
3.19.1 A3						
CREDENCE (A3) - canagliflozin	86	2200	98	2197	100.0%	0.88 [0.66, 1.16]
Subtotal (95% CI)		**2200**		**2197**	**100.0%**	**0.88 [0.66, 1.16]**
Total events	86		98			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.91 (P = 0.36)						
Total (95% CI)		**2200**		**2197**	**100.0%**	**0.88 [0.66, 1.16]**
Total events	86		98			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.91 (P = 0.36)						
Test for subgroup differences: Not applicable						



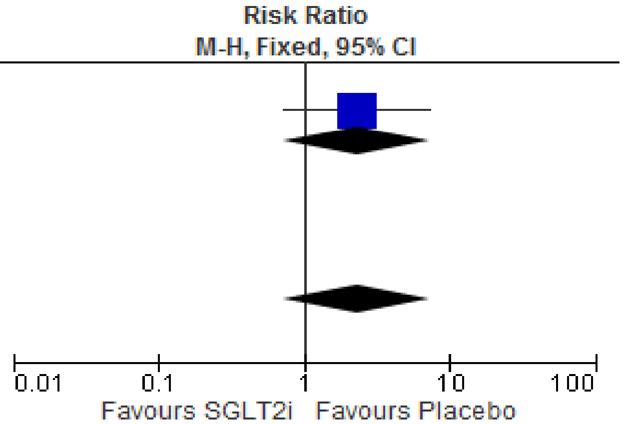
Hypoglycaemia

Study or Subgroup	SGLT2i Events	SGLT2i Total	Placebo Events	Placebo Total	Weight	Risk Ratio M-H, Fixed, 95% CI
3.20.1 A3						
DELIGHT (A3) - dapagliflozin	35	145	29	148	100.0%	1.23 [0.80, 1.90]
Subtotal (95% CI)		**145**		**148**	**100.0%**	**1.23 [0.80, 1.90]**
Total events	35		29			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.94 (P = 0.35)						
Total (95% CI)		**145**		**148**	**100.0%**	**1.23 [0.80, 1.90]**
Total events	35		29			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.94 (P = 0.35)						
Test for subgroup differences: Not applicable						



Genitourinary infection

Study or Subgroup	SGLT2i Events	SGLT2i Total	Placebo Events	Placebo Total	Weight	Risk Ratio M-H, Fixed, 95% CI
3.21.1 A3						
DELIGHT (A3) - dapagliflozin (1)	9	145	4	148	100.0%	2.30 [0.72, 7.29]
Subtotal (95% CI)		**145**		**148**	**100.0%**	**2.30 [0.72, 7.29]**
Total events	9		4			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.41 (P = 0.16)						
Total (95% CI)		**145**		**148**	**100.0%**	**2.30 [0.72, 7.29]**
Total events	9		4			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.41 (P = 0.16)						
Test for subgroup differences: Not applicable						



Footnotes
(1) Combines events for urinary tract infection and genital infection

Appendix F – GRADE tables

SGLT2 Vs placebo

Quality assessment							No of patients		Effect		Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	SGLT2 inhibitor	Placebo	HR/MD/RR (95% CI)	Absolute	
Renal composite – End stage kidney disease, doubling serum creatinine, renal death											
3	Randomised trials	No serious risk of bias	no serious inconsistency	No serious indirectness[2]	no serious imprecision	Composite outcome[2]	4511	3736	HR 0.71 (0.59-0.85)		Moderate
Cardiovascular composite: 3-point MACE											
4	Randomised trials	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	Composite outcome[2]	5040	4485	HR 0.81 (0.73-0.91)		Moderate
All cause mortality											
3	Randomised trials	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	none	4414	4402	HR 0.80 (0.69-0.93)		High
Cardiovascular death											
4	Randomised trials	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	none	5434	5422	HR 0.83 (0.71-0.97)		High
Non-fatal myocardial infarction											
1	Randomised trials	No serious	N/A[4]	No serious indirectness	Serious imprecision[3]	none	2202	2199	HR 0.81 (0.59-1.11)		Moderate

		risk of bias									
Non-fatal stroke											
1	Randomised trials	No serious risk of bias	N/A[4]	No serious indirectness	Serious imprecision[3]	none	2202	2199	HR 0.80 (0.56-1.15)		Moderate
Fatal/non-fatal myocardial infarction											
2	Randomised trials	No serious risk of bias	No serious inconsistency	No serious indirectness	Serious imprecision[3]	none	2232	1627	HR 0.78 (0.59 – 1.02)		Moderate
Fatal/non-fatal stroke											
2	Randomised trials	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	none	2232	1627	HR 0.70 (0.49-0.98)		High
Hospitalisation for heart failure											
3	Randomised trials	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	none	3979	3971	HR 0.58 (0.48-0.71)		High
End stage kidney disease											
2	Randomised trials	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	none	3657	3650	HR 0.68 (0.57-0.82)		High
Doubling of serum creatinine											
1	Randomised trials	No serious risk of bias	N/A[4]	No serious indirectness	No serious imprecision	none	2202	2199	HR 0.60 (0.48-0.75)		High
eGFR reduction >50%											

1	Randomised trials	No serious risk of bias	N/A[4]	No serious indirectness	No serious imprecision	none	1455	1451	HR 0.53 (0.42 – 0.67)		High
Dialysis											
2	Randomised trials	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	none	3657	3653	HR 0.72 (0.57-0.90)		High
eGFR at 6 months											
4	Randomised trials	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision (Estimated MID 3.16) [6]	none	3126	2986	MD -1.91 (-2.83, -0.99)		High
eGFR last available data point >2 years											
3	Randomised trials	No serious risk of bias	Serious inconsistency[5]	No serious indirectness	No serious imprecision (Estimated MID 3)[7]	none	790	581	MD 2.37 (1.75-2.98)		Moderate
Percentage change from baseline UACR 6 months (%)											
1	Randomised trials	No serious risk of bias	N/A[4]	No serious indirectness	No serious imprecision (estimated MID (51.6)[8]	none	148	144	MD -1.00 (-24.84, 22.84)		High
Percentage change from baseline UACR last available data point >2years (%)											
Randomised trials	No serious risk of bias	No serious risk of bias	N/A[4]	No serious indirectness	No serious imprecision (estimated MID 113.7) [9]	none	276	115	MD -38.00 (-81.24, 5.24)		High
Diabetic Ketoacidosis- Canagliflozin											
1	Randomised trials	No serious risk of bias	NA[4]	No serious indirectness	No serious imprecision	none	11/2200	1/2197	RR 10.98 (1.42 – 85.01)	0 more per 1000 (0 more to 0 more)	High
Diabetic Ketoacidosis- Dapagliflozin											

3	Randomised trials	No serious risk of bias	No serious inconsistency	No serious indirectness	Very serious imprecision[10]	none	1/2455	2/2457	RR 0.67 (0.11 - 4.00)	0 more per 1000 (1 fewer to 3 more)	Low
Diabetic Ketoacidosis - SGLT2 class (Canagliflozin or Dapagliflozin)											
4	Randomised trials	No serious risk of bias	Serious inconsistency[5]	No serious indirectness	Very serious imprecision[10]	None	12/4655	3/4654	RR 2.27 (0.21-24.73)	1 more per 1000 (1 fewer to 24 more)	Very low
Amputation - Canagliflozin											
2	Randomised trials	No serious risk of bias	Very serious inconsistency[11]	No serious indirectness	Very serious imprecision[10]	None	104/3220	79/3217	RR 1.48 (0.70-3.13)	12 more per 1000 (7 fewer to 53 more)	Very low
Amputation - Dapagliflozin											
2	Randomised trials	No serious risk of bias	No serious inconsistency	No serious indirectness	Very serious imprecision[10]	none	36/2297	38/2294	RR 0.94 (0.60-1.48)	1 fewer per 1000 (7 fewer to 8 more)	Low
Amputation – SGLT2 class (Canagliflozin or Dapagliflozin)											
4	Randomised trials	No serious risk of bias	Serious inconsistency[5]	No serious indirectness	Serious imprecision[12]	none	140/5517	117/5511	RR 1.28 (0.81-2.02)	6 more per 1000 (4 fewer to 21 more)	Low
Fracture											
5	Randomised trials	No serious risk of bias	No serious inconsistency	No serious indirectness	Serious imprecision[12]	none	194/5674	172/5675	RR 1.13 (0.92 – 1.38)	4 more per 1000 (2 fewer to 11 more)	Moderate
Acute Kidney Injury											
2	Randomised trials	No serious risk of bias	No serious inconsistency	No serious indirectness	Serious imprecision[12]	None	96/3220	113/3217	RR 0.85 (0.65-1.11)	5 fewer per 1000 (12 fewer to 4 more)	Moderate
Hypoglycaemia											

4	Randomised trials	No serious risk of bias	Serious inconsistency[5]	No serious indirectness	Serious imprecision[12]	None	101/2608	142/2771	RR 0.91 (0.72-1.15)	4 fewer per 1000 (12 fewer to 6 more)	Low
Genitourinary infection											
4	Randomised trials	No serious risk of bias	Serious inconsistency[5]	No serious indirectness	Serious imprecision[12]	None	356/2908	282/2750	RR 1.18 (1.02-1.37)	19 more per 1000 (2-38 more)	Low

2. Downgraded due to differences between outcomes used to make composite outcome.
3. Downgraded as the 95% confidence interval crosses the MID (line of no effect).
4. One study included in analysis.
5. I^2 between 33.3% and 66.7%.
6. MID = 0.5 of the median standard deviation of the comparison group (Median SD= 6.3243)
7. MID = 0.5 of the median standard deviation of the comparison group (Median SD= 6.01255)
8. MID = 0.5 of the median standard deviation of the comparison group (SD= 103.2)
9. MID = 0.5 of the median standard deviation of the comparison group (SD= 227.36)
10. 95% confidence interval crosses the MID (0.8-1.25) at both ends
11. I^2 greater than 66.7%
12. 95% confidence interval crosses the MID (0.8-1.25) at one end

SGLT2 vs Placebo (eGFR stratification)

Only stratifications were there was evidence of a difference in effect across stratification levels (test for subgroup differences, p<0.05) are reported.

Quality assessment							No of patients		Effect		Quality
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**SGLT2 inhibitor**	**Placebo**	**HR/MD/RR (95% CI)**	**Absolute**	
eGFR at 6 months eGFR 30-45											

1	randomised trials	no serious risk of bias	N/A[1]	No serious indirectness	Serious imprecision (estimated MID 2.4)[2]	None	228	249	MD -2.30 (-3.02, -1.58)		Moderate
eGFR at 6 months eGFR 45-60											
2	randomised trials	no serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision (estimated MID 8.5)	None	888	761	MD -4.00 (-4.22, -3.77)		High

1. One study included in analysis.
2. Crosses the MID = 0.5 of the median standard deviation of the comparison group (SD= 4.807)

SGLT2 vs Placebo (UACR stratification)

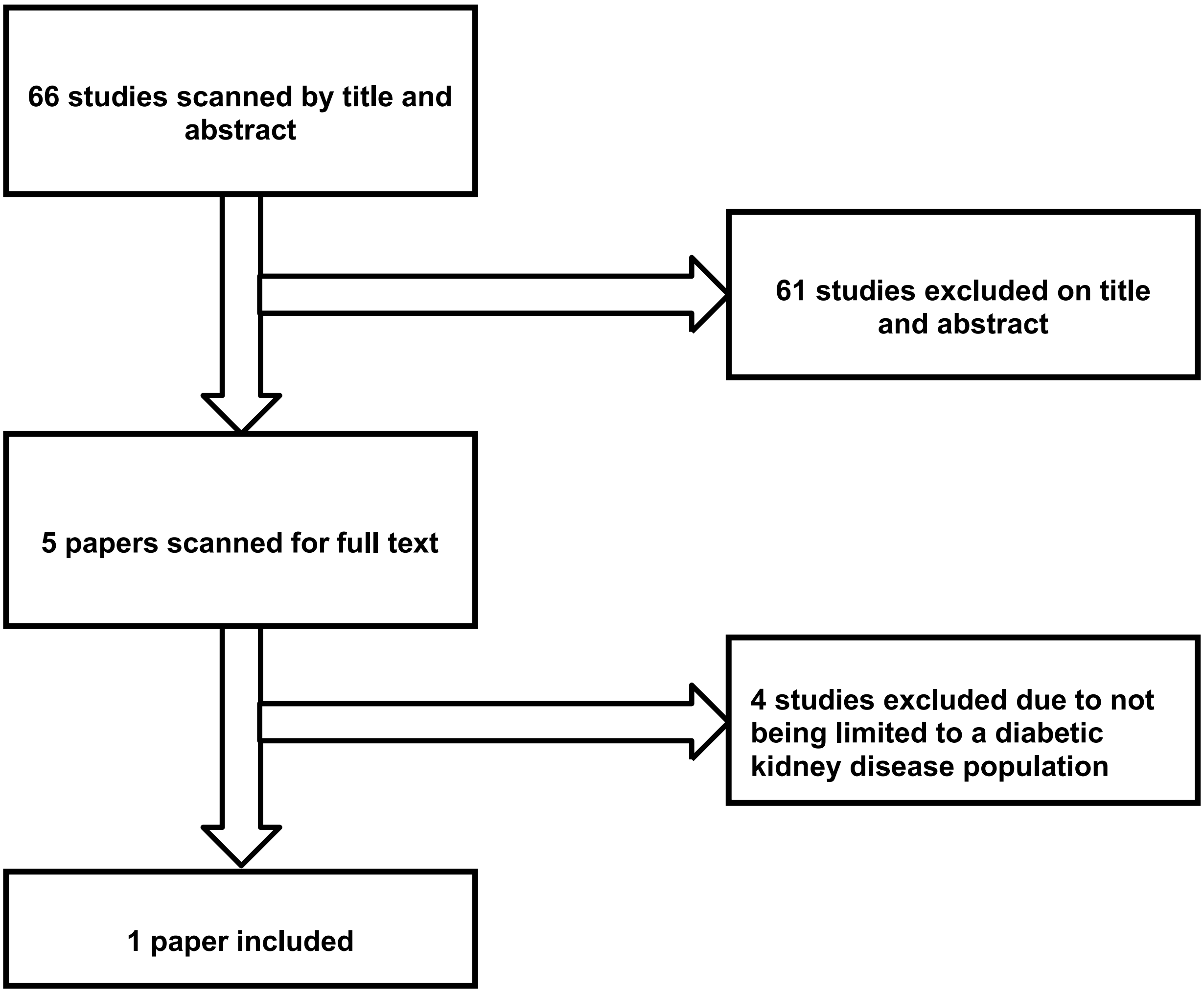
Only stratifications were there was evidence of a difference in effect across stratification levels (test for subgroup differences, p<0.05) are reported.

Quality assessment							No of patients		Effect		Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	SGLT2 inhibitor	Placebo	HR/MD/ RR (95% CI)	Absolute	
eGFR at 6 months UACR A2											
1	randomised trials	no serious risk of bias	N/A[4]	No serious indirectness	No serious imprecision (Estimated MID 3.77)	None	1232	876	MD -1 (-1.71, -0.29)		High
eGFR at 6 months UACR A3											
1	randomised trials	no serious	N/A[4]	No serious indirectness	Serious imprecision[12]	None	373	325	MD -3 (-3.70, -2.30)		Moderate

		risk of bias			(Estimated MID 2.29)						
eGFR at last available data point >2years UACR A2											
1	randomised trials randomised trials	no serious risk of bias	N/A[4]	No serious indirectness	Serious Imprecision[13] (Estimated MID 3.19)	None	349	159	MD 3 (1.89-4.11)		Moderate
eGFR at last available data point >2years UACR A3											
1	randomised trials	no serious risk of bias	N/A[4]	No serious indirectness	No serious imprecision (Estimated MID 2.48)	None	63	26	MD 10 (7.86-12.14)		High

4. One study included in analysis.
12. Crosses the MID = 0.5 of the median standard deviation of the comparison group (SD= 4.518)
13. Crosses the MID = 0.5 of the median standard deviation of the comparison group (SD= 6.3843)

Appendix G – Economic evidence study selection



Appendix H – Economic evidence tables

Table H. 1: Summary of Willis et al. (2021)

Willis et al. (2021). Cost-Effectiveness of Canagliflozin Added to Standard of Care for Treating Diabetic Kidney Disease (DKD) in Patients with Type 2 Diabetes Mellitus (T2DM) in England: Estimates Using the CREDEM-DKD Model	
Study details	**Analysis:** Cost-utility analysis **Approach to analysis:** A discrete event simulation model with an adjustable time horizon. Simulated patients are characterised to be representative of patients in the CREDENCE trial. Baseline characteristics, patient history and time-varying risk factors are used to determine renal health state. The renal health states are supplemented with additional health states relating to MI, stroke, hospitalisation for heart failure and death. DKD progression is experienced in terms of eGFR decline and uACR increase, **Complications considered:** CKD stages 1, 2, 3A, 3B, 4, and 5 prior to dialysis, receiving dialysis, and post renal transplant. Also has health states relating to MI, stroke, hospitalisation for heart failure and death. **Perspective:** United Kingdom **Time horizon:** 10 years **Discounting:** 3.5%
Interventions	**Intervention 1:** Canagliflozin 100mg + SoC **Intervention 2:** Current standard of care (SoC)
Population	**Population:** Adults age 30+ with Type 2 diabetes and CKD, defined as: eGFR: 30 to 90 mL/min per 1.73 m^2 and uACR > 30 mg/mmol **Characteristics:** Mean age: 63; Mean diabetes duration: 15.8; Female: 33.9%
Data sources	**Baseline/natural history:** From CREDENCE population **Risk equations:** Extrapolations of eGFR and log(uACR) were informed by linear mixed model risk equations estimated from CREDENCE, with a minimum threshold of eGFR set at which all patients would immediately assigned to start dialysis. Risk equations were estimated from CREDENCE, and use baseline characteristics, patient history, eGFR and log(uACR) to predict events. **Effectiveness:** From CREDENCE **Costs:** Cardiovascular complications taken from Alva et al. (2015); dialysis and kidney transplant taken from Kerr et al. (2012); CKD stages taken from NICE technology appraisal of tolvaptan for treating autosomal dominant polycystic kidney disease for dialysis and kidney transplant costs [TA358]. **QoL:** Uses a range of sources sourced from a targeted literature search.
Base-case results	
Sensitivity analyses	**Deterministic:** Eight sensitivity analyses performed to check the robustness of the model including varying time horizon, treatment effects on renal, CV, and mortality outcomes, including treatment effects for stroke, dialysis and mortality, removal of eGFR fail-safe floor, and assuming same trajectory of eGFR for both arms. **Probabilistic:** Model structure is patient-level, capturing first and second order uncertainty.
Comments	**Source of funding:** This study was financed by Mundipharma and the fees for the journal’s Rapid Service was supported by Napp Pharmaceuticals Limited (part of the Mundipharma Network)

Willis et al. (2021). Cost-Effectiveness of Canagliflozin Added to Standard of Care for Treating Diabetic Kidney Disease (DKD) in Patients with Type 2 Diabetes Mellitus (T2DM) in England: Estimates Using the CREDEM-DKD Model	
	Limitations: Minor limitations with one potentially serious limitation (see Appendix H)

Table H. 2: Acceptability checklist for Willis et al. (2021)

1.1 Is the study population appropriate for the review question?	Partly (limited to adults over 30)
1.2 Are the interventions appropriate for the review question?	Partly (all available interventions are not included)
1.3 Is the system in which the study was conducted sufficiently similar to the current UK context?	Unclear (Paper states that they have been sourced from a literature review, but elicitation methods are not mentioned)
1.4 Is the perspective for costs appropriate for the review question?	Yes
1.5 Is the perspective for outcomes appropriate for the review question?	Yes
1.6 Are all future costs and outcomes discounted appropriately?	Yes
1.7 Are QALYs, derived using NICE's preferred methods, or an appropriate social care-related equivalent used as an outcome?	Unclear (Paper states that they have been sourced from a literature review, but elicitation methods are not mentioned)
1.8 Overall judgement	**Partially applicable**

Table H. 3: Limitations checklist for Willis et al. (2021)

2.1 Does the model structure adequately reflect the nature of the topic under evaluation?	Yes
2.2 Is the time horizon sufficiently long to reflect all important differences in costs and outcomes?	Partly (the time horizon of 10 years may not be sufficient to capture the lifetime of the patient
2.3 Are all important and relevant outcomes included?	Yes
2.4 Are the estimates of baseline outcomes from the best available source?	Partly (sourced from various sources but not based on literature review, uncertainty about baseline eGFR and uACR progression)
2.5 Are the estimates of relative intervention effects from the best available source?	Partly (from a single RCT instead of a meta-analysis)
2.6 Are all important and relevant costs included	Yes
2.7 Are the estimates of resource use from the best available source?	Yes
2.8 Are the unit costs of resources from the best available source?	Yes
2.9 Is an appropriate incremental analysis presented or can it be calculated from the data?	Yes
2.10 Are all important parameters whose values are uncertain subjected to appropriate sensitivity analysis?	Yes
2.11 Has no potential financial conflict of interest been declared	No

2.12 Overall assessment	**Minor limitations with one potentially serious limitation relating to eGFR/uACR progression**

Appendix I – Health economic model

I.1 Background

As outlined in Section 1.1.7, a review of the economic evidence identified one cost-utility analysis (Willis et al. 2021) that was partially applicable to the review question. Although the study provided some directly applicable evidence, there were a number of potential limitations to using the published results of the study to support the review question:

- Use of a 10-year model time horizon
- Sourcing of baseline characteristics and treatment effects from a single trial (CREDENCE)
- Questions around the clinical plausibility of the extrapolation of dialysis and transplant events
- The exclusion of downstream dialysis costs (an approach which is not aligned to the NICE reference case but is commonly used for modelling dialysis in NICE guidelines; see section I.3.4.2 for further rationale).

The NICE development team noted that these potential limitations were all associated with changeable model inputs rather than being limitations in the model structure. As such, the NICE development team considered that issues around applicability and potential limitations were related to the published results of the study, rather than with the model used in the study.

On this basis, the model used in the Willis et al. (2021) study was adapted to support the review question. The NICE development team would like to thank Dr Willis and the team at The Swedish Institute for Health Economics for giving access to the model, and for sharing their time and expertise. Janssen Global Services, LLC provided access to patient-level CREDENCE study data and financed development of the original model used in the Willis et al. (2021) study and Mundipharma Limited provided financial support for a modelling upgrade. The model was then updated for use in this guideline by the NICE development team.

The adapted model was used to support the committee's consideration of the cost effectiveness of SGLT2 inhibitors as a class for adults with chronic kidney disease and type 2 diabetes. The model was updated to include evidence on treatment effectiveness from the clinical review. Other model inputs (such as cost and quality of life parameters) were updated to use sources that were best aligned to the NICE reference case, and where applicable, values used in NICE guidelines of related disease areas (such as the NICE guideline on Chronic Kidney Disease).

I.2 Model overview

I.2.1 Decision problem

The population considered in the review question was children, young people and adults with chronic kidney disease and type 2 diabetes. The trials identified in the clinical review were all conducted in adults, and so the economic analysis was limited to the adult population with chronic kidney disease and type 2 diabetes. The population was modelled using the baseline characteristics from the CREDENCE trials (see section I.3.1) which was conducted in people with category A3 chronic kidney disease and type 2 diabetes. An exploratory sensitivity analysis also considered the population with category A2 chronic kidney disease.

The interventions included in the model were SGLT2 inhibitors currently licensed for use in people with type 2 diabetes, with doses aligned to the specifications in the summary of product characteristics of each drug. Details of the modelled doses are included in Table I.1. Interventions were modelled as additions to standard care. In line with the scope of this work, the SGLT2 inhibitors were considered as a class; treatment costs were assumed to be equally weighted, and effects were assumed to be class-level.

Table I. 1: Interventions included in the model

Drug	Dose in Summary of Product Characteristics (eGFR as mL/min/1.73m^2)
Canagliflozin	100mg, do not initiate in patients with eGFR <30
Dapagliflozin	10mg, do not initiate in patients with eGFR <60, discontinue if eGFR <45
Empagliflozin	10mg, do not initiate in patients with eGFR <60, discontinue if eGFR <45
Ertugliflozin	5/15mg, do not initiate in patients with eGFR <60, discontinue if eGFR <45

The comparators included in the model were a combination of treatments representing standard care. The committee considered that the standard care treatments included by Willis et al. were representative of current UK practice and so the distributions were not updated. The breakdown of standard care treatments is detailed in Table 2 of the Electronic Supplementary Material 2 in the Willis et al. (2020) study.

In line with the NICE reference case, the economic evaluation was a cost-utility analysis which used a NHS/PSS perspective and applied a 3.5% discount rate to costs and benefits.

A 10-year time horizon was used in the base-case analysis. The committee considered that there was considerable uncertainty in the extrapolation of eGFR and uACR past the trial period (see section I.3.2.1) and so opted to explore scenarios with reduced time horizons (see section I.4). The NICE reference case that states that that the time horizon should be long enough to reflect all important difference in costs or outcomes between the technologies being compared. However, restricting the time horizon to the trial follow-up period is unlikely to meet this criterion, and means that the costs of treatments are incurred without downstream benefits of treatments being captured. Therefore, a 40-year time horizon was also modelled in a sensitivity analysis .

I.2.2 Model structure

A full description of the model structure is outlined in Willis et al. (2021), with a description of the development and internal validation of the model outlined in Willis et al. (2020). A schematic of the model, taken from the Willis et al. (2021) paper, is outlined in Figure I.1. In brief, a cohort of patients is simulated at baseline with values for eGFR, log(uACR), age, sex, smoking, diabetes duration and history of cardiovascular event (MI, stroke or hospitalisation for heart failure). The values for eGFR and log(uACR) for each patient are extrapolated over time, in line with a model estimated by Perkovic et al. (2019). The model runs in 26-week cycles; for every cycle, the updated eGFR and log(uACR), patient history and characteristics are used to predict the CKD stage a patient is in, whether they have had a cardiovascular event, and indirectly, whether they have dialysis or a transplant.

The CKD stages included in the model are Stage 1, Stage 2, Stage 3a, Stage 3b, Stage 4 and Stage 5 pre- renal replacement therapy. Patients in Stages 3b to Stage 5 have the potential to move to dialysis or transplant, which are captured in ongoing dialysis and post-transplant states. Movement to the dialysis state is estimated via a risk-factor equation which accounts for eGFR, log(uACR), history of events and patient characteristics. Further information about the selection of risk-factor equations used in the model are detailed in Willis et al. (2020). There were too few transplant events observed in the CREDENCE trial to estimate a risk-factor equation, so transition to the transplant state from Stages 3b to Stage 5 follows a user-defined probability. The model also allows for a user-defined eGFR threshold below which a patient will start either dialysis or have a transplant (the probability of receiving one type of renal replacement therapy over the other is user-defined).

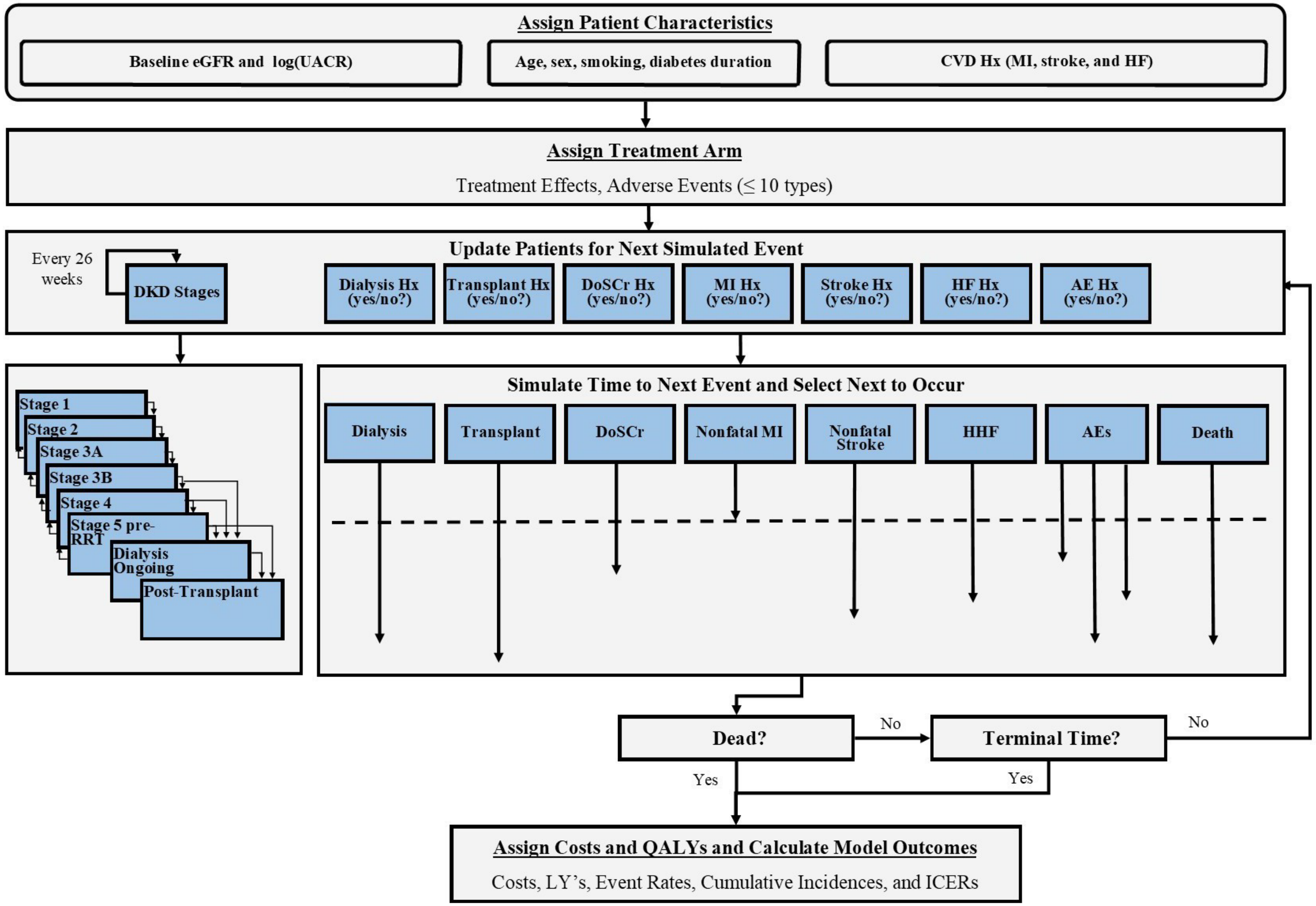
The cardiovascular events included in the model are non-fatal MI, non-fatal stroke, hospitalisation due to heart failure and death. These are also estimated via risk-factor equations based on eGFR, log(uACR), history of events and patient characteristics; further details about equation selection are outlined in Willis et al. (2020).

Treatment effects can be applied via the eGFR and log(uACR) slopes to affect the trajectory of these extrapolations, which in turn affects the prediction of CKD stages, renal replacement therapy and cardiovascular events. Hazard ratios can also be applied directly to cardiovascular events, mortality and start of renal replacement therapy. Rates and duration of additional adverse events can also be input for the intervention arm.

Costs and utility values are ascribed to the CKD stages, cardiovascular and renal events and adverse events for each patient, to enable to the model to output ICERs per QALY gained for the SGLT2 inhibitors + standard care compared to standard care alone.

The model is a patient-level simulation which requires the user to define the number of patients in the cohort and number of simulations. All analyses were run for a cohort of 500 patients over 500 model runs to achieve model convergence in the original analysis. This led to an extended model time.

Figure I. 1: Model schematic



DKD diabetic kidney disease, DoSCr doubling of serum creatinine, hx history, HF heart failure, HHF hospitalization for heart failure, ICER incremental cost-effectiveness ratios, LY life year, MI myocardial infarction

I.3 Model inputs

I.3.1 Baseline characteristics

The population comprised adults with DKD and type 2 diabetes mellitus (T2DM), with baseline characteristics based on the patients enrolled in the CREDENCE trial (Table I.2), consistent with the modelled population in the Willis et al. (2021) analysis. In the base case population, patients were, at least 30 years of age, had estimated glomerular filtration rate (eGFR) between 30 to 90 mL/min/1.73 m^2, and urinary albumin to creatinine ratio (uACR) greater than 30 mg/mmol. Patients received a maximum tolerated dose of either an angiotensin-converting enzyme (ACE) inhibitor or angiotensin receptor blocker (ARB).

In order to capture correlation between characteristics, individual patient data from CREDENCE was analysed for the model published by Willis et al. (2021). Since these patient characteristics are not independent of each other, values were sampled with correlation defined by variance-covariance matrix to support risk factor clustering.

The possibility of pooling characteristics from each of the DKD populations of the trials in the evidence review was considered. However, only summary data was available and without individual patient data, it is not possible to estimate the covariance between each characteristic. Therefore, the base case model retained the baseline values estimated from CREDENCE in the Willis model.

The committee agreed that these baseline values were representative of the DKD A3 (uACR greater than 30 mg/mmol) population. Some of the trials in the evidence review also included patients with uACR less than 30 mg/mmol and provided analyses that were stratified by this baseline characteristic. Therefore, a subgroup analysis considered a population with uACR between 3 to 30 mg/mmol. The evidence review found that no studies reported outcomes separately for the A1 population (uACR less than 3 mg/mmol) and so it was not possible to conduct any analyses for this subgroup nor to explore it in the economic analysis.

Table I.2 Baseline characteristics

Characteristic	Mean	Standard deviation
Age (years)	63.0	9.2
Female	33.9%	0.47
Smokers	14.5%	0.35
Diabetes duration (years)	15.8	8.6
Baseline eGFR (mL/min)	56.2	18.2
Baseline log uACR	6.8	1.0
Previous history of MI	10.0%	0.30
Previous history of stroke	10.4%	0.31
Previous history of heart failure	14.8%	0.36

I.3.2 Treatment effects

I.3.2.1 eGFR and uACR

Both eGFR and log(uACR) determine which CKD stage a patient falls into and feed into the risk-factor equations predicting cardiovascular and renal events.

The model extrapolates log(uACR) rather than uACR. Log(uACR) is not reported as an outcome in the trials included in the clinical review and cannot be derived without patient level data on ACR. It was not feasible to obtain patient level trial data for relevant SGLT2 trials within the development period and so an a priori judgement was made to retain the extrapolations of log(uACR) derived from CREDENCE in the Willis et al. (2021) analysis. The values used in the Willis et al. (2021) analysis are outlined in Table I.3. The Willis et al. (2021) and (2020) studies both refer to the Perkovic et al. (2019) study for the original derivation of the log(uACR) extrapolation function, however the reporting of the methods used is somewhat opaque. This was recognised by committee as a limitation of the model and a source of uncertainty in the resulting cost-effectiveness estimates.

Table I. 3: Default In(uACR) progression in model for intervention

uACR progression (proportional change)	Mean	95% CI (lower)	95% CI (upper)
Initial effect	-0.31	-0.36	-0.26

The extrapolation of eGFR in the model is based on the analysis by Perkovic et al. (2019). This fitted a non-parametric curve to the eGFR data observed in CREDENCE to estimate a model with two linear slopes that has a ‘knot’ at week 3 where the hazard changes from one linear slope to the other. A treatment effect can be incorporated by altering the gradient of the pre- and post- 3-week slopes for the intervention arm. The values used in the Willis et al. (2021) analysis are outlined in Table I.4, which provides the relative difference in change in eGFR between SGLT2 inhibitors and standard care. The negative value for the initial effect indicates that patients on SGLT2 inhibitors experience a larger decline in eGFR than patients on standard care up to three weeks, and the positive value for the slope per year indicates that subsequently they experience a smaller decline in eGFR.

Table I. 4: Default eGFR progression in model for intervention

eGFR progression (absolute)	Mean	95% CI (lower)	95% CI (upper)
Initial effect (pre- 3-weeks)	-3.2	-3.9	-2.5
Slope p/year (post- 3-weeks)	2.74	2.37	3.11

As with log(uACR), the reporting of the methods used to derive the extrapolation function are fairly opaque. The clinical review extracted trial data on mean change in eGFR at 6 months and at the last available measurement, and so the NICE development team considered whether this information could be used to derive an alternative extrapolation of eGFR which incorporated data on multiple SGLT2 inhibitors. However, in the absence of patient level data any alternative approach to extrapolation was likely to have substantial methodological limitations. There was also some uncertainty associated with the estimates of eGFR extracted from the clinical review as in some instances they were not reported directly and were instead extracted from graphs, meaning there was a potential margin of error in the estimates. For these reasons, the original extrapolations of eGFR derived from CREDENCE in the Willis et al. (2021) analysis were retained. Again, the committee recognised this as a limitation of the model and a source of uncertainty in the resulting cost-effectiveness estimates.

I.3.2.2 Start of dialysis and transplant

The model includes three mechanisms by which a treatment effect can be applied to the start of dialysis and transplant events:

1. Via eGFR and log(uACR):
    a. eGFR and log(uACR) affect progression through the CKD stages. A risk-factor equation (dialysis) or fixed probability (transplant) is then used to estimate progression from CKD Stages 3b to Stage 5 to either form of renal replacement therapy.
    b. An eGFR threshold below which patients are assumed to automatically start either dialysis or transplant, and a parameter determining the probability of each mode.
2. An option for a hazard ratio to be applied directly to the transition rates or dialysis and transplant events.

The risk-factor equation for dialysis was retained from the original model in the absence of the patient level data needed to update the equation. In the Willis et al. (2021) analysis the probability of transplant is assumed to be 0 for Stages 3b and 4 on the premise that there were no transplant events observed for patients in these stages in the CREDENCE trial. Data on the timing of kidney transplant was not extracted from the clinical review, and so this assumption was retained in the updated analysis.

Values for the probability of transplant in Stage 5 pre- renal replacement therapy (RRT) and Stage 5 on dialysis were taken from the UK Renal Registry 22nd Annual Report. Details of the values used are outlined in Table I.5. In the case of those patients with both CKD and T2D, the probability of having a transplant is most likely lower in comparison to the overall CKD population, since patients with T2D are often ineligible for transplantation due to their numerous comorbidities. We are not aware of any data on the number of kidney transplants specific to diabetic patients, and so the values originally applied in the Willis et al. (2021) analysis were explored in a scenario analysis.

Table I. 5: Transplant rates used in the model

Stage	Rate p/year	Source
Stage 3b	0.000	Original assumption from Willis et al. (2021)
Stage 4	0.000	
Stage 5 pre-RRT	0.081	UK Renal Registry 22nd Annual Report Table 1.6
Stage 5 on dialysis	0.125	UK Renal Registry 22nd Annual Report Table 3.3 (new transplants [2018]) Tables 4.2, 5.2 and 6.2 (number of people on dialysis in previous year [2017])

In the Willis et al. (2021) analysis the eGFR threshold at which a patient would start renal replacement therapy was assumed to be an eGFR of 6 ml/min/1.73m^2. All patients meeting this threshold were assumed to start dialysis. The proportion of patients starting renal replacement therapy through this mechanism was low and so had a minimal impact on overall results. For this reason, the updated analysis retained this threshold and distribution of modes.

The CREDENCE trial reported a hazard ratio of 0.74 (0.55 to 1.00) for start of dialysis or kidney transplant. However, as the number of transplant events in CREDENCE was low this hazard ratio was only applied to start of dialysis in the Willis et al. (2021) analysis and the hazard ratio for kidney transplants was set to 1. The clinical review extracted a hazard ratio of 0.72 (0.57 to 0.90) for start of dialysis which was used in the updated analysis. The clinical review did not extract data on kidney transplant events; as none of the trial periods were long enough to capture many transplant events the same assumption of a hazard ratio of 1 was retained.

I.3.2.3 Cardiovascular and mortality outcomes

Treatment effects for cardiovascular and mortality outcomes can be applied using hazard ratios which are applied for the duration of treatment. The model includes events for non-fatal MI, non-fatal stroke, hospitalisation for heart failure and all-cause mortality. The model includes an option to model separate hazard ratios for the first year and subsequent years. There is a further option to model an odds ratio for the cause of death being cardiovascular related rather than non-cardiovascular related.

Treatment effects for hospitalisation for non-fatal MI, non-fatal stroke, hospitalisation for heart failure and all-cause mortality were taken from the clinical review. The clinical review did not extract an odds ratio for the cause of death being cardiovascular related and so the value from the original analysis was used, although there was some opacity in the reporting of this parameter. Full details of the hazard ratios are outlined in Table I.6.

In the original model, hazard ratios for non-fatal MI and hospitalisation for heart failure were applied for the first year and subsequent years, whereas the hazard ratio for non-fatal stroke was only applied after the first year. Willis et al. (2020) suggests that this was because there was limited treatment effect observed in CREDENCE for non-fatal stroke over the first year of the trial (Mahaffey et al. 2019). Trials in the evidence review reported hazard ratios for the whole-trial period and did not provide information on whether or how they changed over time, and so the assumptions made in the original model were retained.

Table I. 6: Hazard ratios applied to SGLT2 inhibitors in the model

Endpoint	Hazard ratio (95% CI)		Source
	First year	Subsequent year	
Non-fatal MI	0.81 (0.59, 1.10)		Mahaffey et al. (2019)
Non-fatal stroke	1.00 (1.00, 1.00)	0.80 (0.56, 1.15)	
Hospitalisation for heart failure	0.58 (0.48, 0.71)		Clinical review
All-cause mortality	0.85 (0.52, 1.37)		

I.3.3 Adverse event rates

The rates of adverse events (AEs) were obtained from the Willis et al (2021) model. AEs were included only for the SGLT2 inhibitor arm, estimated as the difference between canagliflozin and standard care in CREDENCE, and consisted of urinary tract infection, genital mycotic infection, diabetic ketoacidosis, and lower extremity amputation (LEA). The event rates, reflecting all-grade AEs, are presented in Table I.7. Since the model was configured to estimate AEs from the between-arm difference in event rate, it was not possible to use outcomes from the evidence review on AEs as they were not reported in this manner. Similarly, there were a number of additional adverse events that the committee felt were significant, such as osteoporosis, acute kidney injury, hypotension/falls, hypoglycaemia and fractures, that could not be included due to the difference in how they are reported by trials and how they were incorporated into the model framework. The impact of these was considered qualitatively by the committee alongside the cost-effectiveness results, and it was decided that they did not expect omission of these events to change the conclusion of the cost-effectiveness analysis, given that they are a relatively small component of total costs and quality of life impacts.

The committee noted that in other areas of diabetes modelling, hypoglycaemic events could sometimes have a substantial impact on cost-effectiveness results. This is because hypoglycaemic events often happen in the short-term and so are discounted less heavily than future micro- and macrovascular events. However, the committee noted that the clinical review found no statistically significant difference in the rates of hypoglycaemic events in SGLT2 inhibitors and standard care. On this basis, the committee were content to accept the omission of hypoglycaemia from the model.

The cost of managing UTI and GMI was based on the existing values from Willis et al. (2021), which took values from the NICE appraisal of canagliflozin, dapagliflozin and empagliflozin for type 2 diabetes (NICE technology appraisal 390). The remaining AE costs were updated for this guideline update: the cost of DKA was obtained from NHS Reference Costs 2018/2019. Amputation was considered to have a permanent impact on management and was associated with both a one-off event-related cost and an ongoing cost of management, and these costs were obtained from Alva et al. (2015), an analysis of the impact of diabetes-related complications on health care costs, and adjusted for inflation to 2020 values.

Disutility values for each event were based on the existing values from Willis et al. (2021). The disutility of adverse events were assumed to apply for one week. Disutility values for UTI and GMI were obtained from studies of health state values, diabetes-related complications and treatment-related adverse events in type 2 diabetes.

Table I.7 Adverse event rates, disutility and management cost

Event	Annual event rate (additional rate for SGLT2i)	Disutility	AE duration (days)	Event-related cost	Subsequent cost per day
UTI	0.0032	-0.0043 [1]	7	£90.91	£0
GMI (male)	0.0075	-0.0046 [1]	7	£56.06	£0
GMI (female)	0.0065	-0.0046 [1]	7	£52.45	£0
DKA	0.0020	-0.0091 [2]	7	£1561	£0
Amputation	0.0011	-0.1690 [3,4]	7	£14,041	£10.68
References: [1]Shingler et al. (2015). [2]Peasgood et al. (2016). [3]Clarke et al. (2002). [4]Sullivan et al. (2016)					

I.3.4 Costs

The perspective for costs and outcomes was that of the NHS and PSS. Unit costs and resource use were obtained from national sources, the current NICE guideline for CKD and the published literature. A summary of costs used in the cost-effectiveness analysis are provided in Table I.8.

I.3.4.1 Treatment costs

Unit costs of the SGLT2 inhibitors (canagliflozin, dapagliflozin, empagliflozin, ertugliflozin) were obtained from the British National Formulary (BNF) and the posologies from the respective SPCs (see Table I.1). In order to estimate an average cost of SGLT2 inhibitors (£453.84 per year), the analysis assumed equal proportions of patients receiving treatment with each SGLT2 inhibitor; since prescription data does not report usage by indication and SGLT2 inhibitors are also licensed for patients with type 2 diabetes, it was not possible to obtain more accurate estimates of usage. However, three of the SGLT2 inhibitors have relatively similar costs and so this was not thought to be a major limitation. No administration costs were applied to these costs and costs were estimated on the basis of 100% compliance. In the base case analysis, it was assumed that patients would receive treatment until death, and discontinuation at a pre-specified time point corresponding to the estimated time in which the cohort reached the eGFR threshold was considered qualitatively as it was not possible to incorporate this into the model structure.

At the time of the analysis, a confidential Commercial Medicines Unit (CMU) discount was available for canagliflozin, which would apply when prescribed in a secondary care setting. There was uncertainty in the proportion of patients that would be prescribed in each setting; expert opinion from the committee suggested that it may be 50%. As such, the base case analysis took the more conservative position that all patients would be prescribed in primary care (i.e. the discount was not applied) and the impact of applying the discount was explored in scenario analyses.

The cost of standard care was obtained from Willis et al (2021), which calculated costs on the basis of use of background therapy in CREDENCE, separately by arm, with unit costs from the BNF. This was estimated as £259.40 annually for those receiving SGLT2 inhibitors and £261.83 annually for those receiving standard care alone. The most commonly used medications included insulin (65% to 66%), statins (68% to 70%), anti-thrombotics (58% to 61%) and biguanides (58%). Further details of the medications constituting standard care and their rates of use are provided in Willis et al (2021). While the committee provided feedback that there may be some differences in standard care in CREDENCE and in current practice, the impact on cost-effectiveness was predicted to be negligible given their relatively low cost and similar levels of use in both treatment arms.

I.3.4.2 Health state costs

The cost of diabetes-related cardiovascular events, including myocardial infarction (MI), hospitalisation for heart failure (HHF), stroke and CV death included a one-off event-related cost and an ongoing cost of management. These costs updated from the values used by Willis et al (2021) to those used in NICE guideline for type 2 diabetes, and were obtained from Alva et al. (2015) and inflated to 2020 prices using the NHSCII index from PSSRU. This study estimated the immediate and long-term inpatient and non-inpatient costs of type 2 diabetes patients, including consultations, visits, admissions and procedures. Resource use in the study was obtained by looking at inpatient use as obtained from the HES database and non-inpatient costs were obtained using questionnaires.

The long-term costs of dialysis were estimated in line with the approach in the NICE CKD guideline update. This was estimated to be £21,607 per year, using unit costs for each mode of dialysis from NHS Reference Costs and an estimate of the number of sessions per cycle for each type of dialysis from NICE Technology Appraisal 117, weighted according to the mode of administration with usage reported by Renal Registry 22nd Annual Report.

The cost of transplant was an average of costs from living and deceased donors, taken from NHS reference costs 2018/2019. Ongoing costs of immunosuppressive therapy applied with the post-transplant health state were obtained from the NICE CKD guideline. People were assumed to use immediate-release tacrolimus and mycophenolate mofetil, with average doses of 0.2 mg/kg/day for tacrolimus and 2g/day for mycophenolate mofetil taken from an HTA report. Costs for these medications were taken from the NHS drug tariff (September 2020) and weighted by usage data for each product and dose from the NHS Prescription Cost Analysis (March 2019).

The costs for DKD stages were retained from Willis et al (2021) and inflated from 2019/2020 prices using the NHSCII index from PSSRU, which were sourced from a NICE appraisal of tolvaptan for treating autosomal dominant polycystic kidney disease (TA358). These costs reflect the background management of patients with DKD and include resources such as visits with health care professionals (consultant, specialist nurse), and clinical tests (biochemistry test, haematology test and phlebotomy). When considering the results of the cost-effectiveness analyses, the committee reflected that the cost of managing patients in DKD Stage 3 was higher than expected. The cost was estimated for patients with CKD who are typically managed in secondary care at this stage, and the committee considered that patients with DKD in Stage 3 are more likely to be managed in primary care. This was taken into account in their decision making: since SGLT2 inhibitors improved survival and slowed disease progression, patients would spend more time in the DKD Stage 3 health state than those on standard care alone, and so an overestimate of costs in this health state would provide a conservative estimate of cost-effectiveness for SGLT2 inhibitors.

There was some uncertainty whether there was double counting of the costs associated with DKD stage and with CV events, since costs values are attributed to the DKD stage *and* the events they experience. This depends on whether being in a DKD stage inherently has an effect on costs, or whether DKD stages are predictors of events, and it is the events that have an effect on costs. To address this uncertainty, a scenario analysis was undertaken where the costs for DKD stages were removed.

Table I.8 Summary of cost inputs in the economic analysis

Resource	Cost	Source and assumptions
SGLT2 inhibitors	£453.84 per year	BNF Canagliflozin: £39.20 per 30 pack Dapagliflozin: £36.59 per 28 pack Empagliflozin: £36.59 per 28 pack Ertugliflozin: £29.40 per 28 pack
Standard of care	SoC: £261.83 per year SGLT2i: £262.05 per year	Willis et al (2021)
Nonfatal MI event	£8,419	Alva et al (2015)
MI history	£2,093	Alva et al (2015)
Nonfatal stroke	£9,054	Alva et al (2015)
Stroke history	£2,157	Alva et al (2015)
Nonfatal HHF	£4,782	Alva et al (2015)
CHF history	£2,805	Alva et al (2015)
CV death	£3,080	Alva et al (2015)
Other death	£3,080	Alva et al (2015)
Stage 1	£193.31	Willis et al. (2021)
Stage 2	£193.31	Willis et al. (2021)
Stage 3a	£1,615	Willis et al. (2021)
Stage 3b	£1,615	Willis et al. (2021)
Stage 4	£3,776	Willis et al. (2021)

Resource	Cost	Source and assumptions
Stage 5 pre-RRT	£5,892	Willis et al. (2021)
Dialysis	£21,607	NHS Reference Costs (2020)
Transplantation Event	£12,565	NHS Reference Costs (2020)
Post-Transplant	£8,332	NHS drug tariff (Sep 2020) PCA (Mar 2019)

I.3.5 Health-related quality of life

Utility values for DKD health states and DKD-related cardiovascular events are summarised in Table I.9. Health state utilities were estimated by applying a disutility relative to a baseline value of 0.785: for example, the disutility for MI is -0.06 and therefore the health state utility value is calculated as 0.725 (0.785-0.06). The utility and disutility values were retained from the model developed by Willis et al (2021), who identified utility values representative of the UK population with DKD via a targeted literature review.

The utility values used in the Willis model for dialysis and transplant were taken from Lee et al (2005), which provided disutility estimates of -0.53 for dialysis and -0.29 for transplant. These were considered to lack validity as applying these to the baseline utility value would give someone on dialysis a utility value of 0.255, which was considered to be very low for these patients. Therefore, the values used in the model were obtained from Beaudet et al. (2014), a systematic review of utility values in type 2 diabetes. The estimate for dialysis is a weighted average of the haemodialysis and peritoneal dialysis disutilities, weighted by the distributions of modes of administration taken from the UK Renal Registry 22nd annual report. The committee felt that the value for dialysis may be too optimistic and may not have captured the burden of symptoms of a patient in this health state, and that it provided an estimate for quality of life than is better than that of patients in DKD Stage 3. Since SGLT2 inhibitors reduce the rate at which dialysis occurs, using the smaller utility decrements for dialysis represents a conservative scenario with respect to the cost-effectiveness of SGLT2 inhibitors, as less value is assigned to averting these events. To address this uncertainty, the impact of using disutility values from Lee et al (2005) for these two health states were explored in a scenario analysis.

Table I.9 Summary of utility values

Health state	Mean	Source
Health state utility		
Baseline utility	0.785	Clarke et al. (2002)
Health state or event-related disutility		
MI	-0.06	Clarke et al. (2002)
Stroke	-0.16	Clarke et al. (2002)
CHF	-0.11	Clarke et al. (2002)
DKD stage 1	-0.15	Jesky et al. (2016)
DKD stage 2	-0.15	Jesky et al. (2016)
DKD stage 3a	-0.20	Jesky et al. (2016)

Health state	Mean	Source
DKD stage 3b	-0.20	Jesky et al. (2016)
DKD stage 4	-0.26	Jesky et al. (2016)
DKD stage 5 (pre-RRT)	-0.27	Jesky et al. (2016)
Dialysis	-0.176	Beaudet et al. (2014)
Post-transplant	-0.023	Beaudet et al. (2014)

I.4 Scenario analyses

I.4.1 Varying the model time horizon

The base case analysis estimated outcomes over a time horizon of ten years. In order to address uncertainty around extrapolation of eGFR and uACR progression, particularly due to implications for predicted dialysis and kidney transplant rates, two scenarios with a shorter time horizon and one scenario with a longer time horizon were conducted. The scenarios were a) 2.6 years, the median follow-up of the CREDENCE trial, b) 5 years, c) 40 years, capturing the entirely of the remaining lifetime, with 1-2% estimated to be alive at this time point.

I.4.2 Commercial medicines unit (CMU) discount applied to canagliflozin

A confidential CMU discount available for canagliflozin applies when it is prescribed in a secondary care setting. Given the uncertainty in the proportion of patients that would be prescribed in each setting, the base case analysis assumed that the discount would not apply to any patients, providing the most conservative estimate of cost-effectiveness in this regard. Firstly, the CMU discount was applied in the cost-effectiveness analysis under all base case assumptions and input values. Since the addition of a discount to canagliflozin would only make it more cost-effective, further scenarios were conducted where the discount was applied to any analyses in which SGLT2 inhibitors were not cost-effective, to determine whether the addition of the discount would change the conclusions of the analysis.

In these scenarios, it was assumed that 100% of patients on canagliflozin are prescribed in secondary care (i.e. CMU discount is applied to all patients on canagliflozin). This provides a range of the likely cost-effectiveness estimates.

I.4.3 Removal of CKD stage costs and utility values

Cost and disutility values are attributed to both the DKD stage *and* the CV events they experience. As there was uncertainty whether this constituted double counting of the costs and disutility values, a scenario analysis was undertaken where the costs and disutility values for DKD stages were removed.

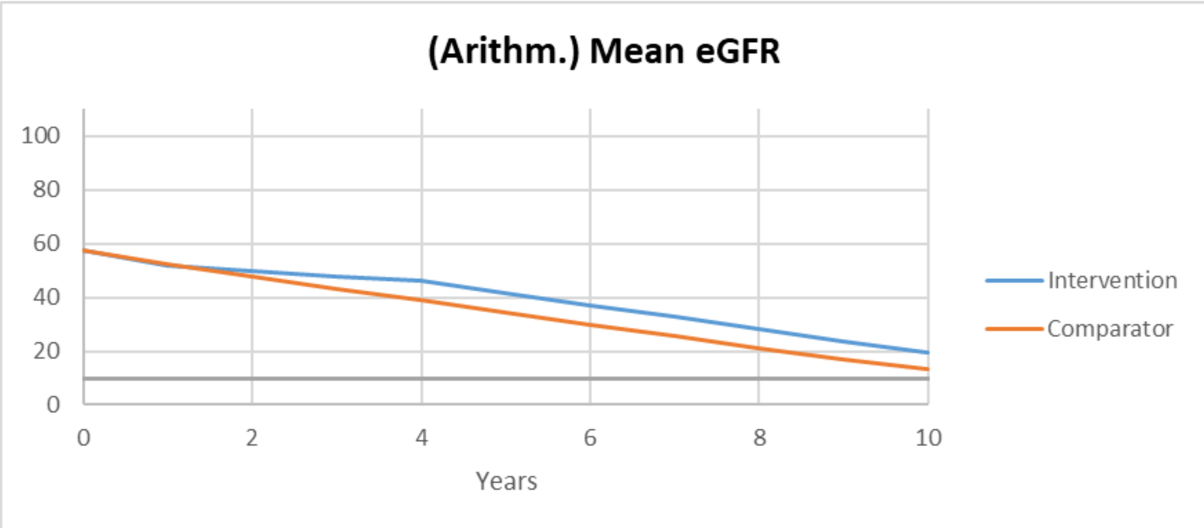
I.4.4 Utility values for dialysis and transplant from Willis et al. (2021)

The cost-effectiveness model took alternate values to those used in the Willis model for dialysis and transplant, as there were concerns that these lacked validity (Section I.3.5). For example, applying the disutility values for dialysis to the baseline utility value would give a utility value of 0.255, which is very low and more consistent with estimates for patients with advanced, progressive cancers. Nevertheless, a scenario analysis explored the impact of applying the disutility values from this source.

I.4.5 Waning of treatment effect after 4 years

Waning of treatment effect was explored to assess the impact to cost-effectiveness in scenarios such as when the treatment effect wanes over time due to causes such as biological resistance to therapy, or fluctuating adherence and disease control over time. This scenario was also conducted to provide indirect cost-effectiveness evidence corresponding to patients discontinuing treatment after a set time or threshold. A time point of four years was selected for the treatment waning effect to occur, as this is when eGFR is predicted to be 45 mL/min/1.73m^2, which is when the license for empagliflozin, dapagliflozin and ertugliflozin states that patients discontinue treatment. Treatment waning was implemented by assuming the rate of eGFR progression and event hazard rate for SGLT2 inhibitors were the same as in the standard care arm. With eGFR progression, this assumes that the initial treatment benefit is maintained over time as illustrated in Figure I.2. Due to the structure of the model, it was not possible to conduct a scenario whereby the absolute eGFR rate of SGLT2 inhibitors converges to that of standard care, which would represent the most conservative scenario.

Figure I.2 eGFR progression over time in the treatment waning scenario



I.4.6 Alternative rates of transplant

Values for the probability of transplant were taken from the UK Renal Registry 22nd Annual Report. However, these may overestimate the rate of transplant in the modelled population as the probability of having a transplant for patients with both CKD and T2D is most likely lower in comparison to the overall CKD population, since patients with T2D are often ineligible for transplantation due to their numerous comorbidities. This scenario analysis explored the impact of using the rates originally applied in the Willis et al. (2021) analysis which were lower than those used in the base case analysis, and were 0.075 (0.053 to 0.097) per year.

I.5 Subgroup analysis

The base case population was based on evidence and baseline characteristics of patients with uACR of 30mg/mmol or more (A3 category). An exploratory subgroup analysis considered a population with less severe proteinuria, with uACR between 3 and 30 mg/mmol (A2 category).

In this analysis, subgroup-specific effects on outcomes derived from the clinical review were applied to the analysis, where available. The only outcomes available for the A2 population in the clinical review were all-cause mortality and heart failure. As illustrated in Table I.10, the HR reported for outcomes in A2 are very similar to those in A3. There was no evidence in this population for the treatment effect on the other outcomes in the model. The meta-analysis of eGFR at 6 months and at one year suggested an ACR subgroup effect on eGFR; however, it was not possible to include this evidence in the model due to limitations described earlier. As such, eGFR progression in this analysis (along with the other outcomes not listed in Table I.10) were based on evidence for the A3 population: this was considered a major limitation of this subgroup analysis.

Table I.10 Treatment effect in the uACR 3 to 30 mg/mmol population

Outcome	Hazard ratio - A2 population	Hazard ratio - A3 population (base case)
Chronic heart failure	0.57 (95% CI 0.34 to 0.95)	0.58 (95% CI 0.48 to 0.71)
All-cause mortality	0.78 (95% CI 0.55 to 1.11)	0.80 (95% CI 0.69 to 0.93)

There was no evidence from trials in the review that provided an estimate of baseline characteristics in this subgroup. uACR was thought not to be normally distributed, and so the mean value was assumed to be the midpoint of the range on a log scale with a minimum value of 3 and a maximum value of 30, and was calculated as approximately 9.4. The remainder of the patient baseline characteristics remained the same as in the base case analysis. While the clinical review stratified the analyses by baseline uACR and eGFR to look at treatment effect, it did not look at the relationships between baseline uACR and eGFR or other baseline characteristics. Without individual patient data from these trials, it was not possible to capture correlations between uACR and other baseline characteristics.

I.6 Results

I.6.1 Base case analysis

Table I.11 presents the results of the base case analysis in population of patients with uACR > 30 mg/mmol. These analyses capture both first-order (population variation) and second-order (parameter uncertainty) uncertainty, and are estimated from 500 cohorts each of 500 patients generated by the economic model. Total costs and QALYs were estimated at ten years, with future outcomes discounted at 3.5%.

The analysis estimated that SGLT2 inhibitors in addition to standard care are a dominant treatment option compared with standard care alone, being both cost saving and producing more benefits. The acquisition cost of SGLT2 inhibitors contributes to the largest gain in costs (see Table I.12 for a breakdown of total costs), and there are also some additional costs attributed to management in CKD stage 3a. However, these additional costs are offset by cost savings due to averted transplants and fewer patients in the more severe and costly to manage CKD stages.

Table I.11 Base case results

	Total discounted costs	Total discounted QALYs	Incremental costs (95% CI)	Incremental QALYs (95% CI)	ICER (£ per QALY)
SGLT2 inhibitors + standard care	£35,467	3.90	-£3,709 (£5,444, -£2,095)	0.24 (0.18, 0.3)	Intervention dominates
Standard care	£39,177	3.66	-	-	-

Table I.12 Breakdown of total costs at ten years

Cost	SGLT2 inhibitors + standard care	Standard care	Incremental costs (95% CI)
SGLT2 inhibitors	£3,167	£0	£3,167 (£2,745, £3,346)
Standard care	£1,884	£1,802	£82 (£54.88, £110.46)
CKD stage 2	£249	£177	£72 (£40, £108)
CKD stage 3a	£3,647	£2,476	£1,171 (£888, £1,497)
CKD stage 3b	£3,916	£3,238	£678 (£222, £1,060)
CKD stage 4	£2,907	£5,293	-£2,386 (-£3,388, -£1,351)
CKD stage 5 (pre-RRT)	£143	£2,437	-£2,294 (-£3,067, -£1,241)
Dialysis	£8,794	£11,027	-£2,233 (-£3,803, -£969)
Transplant	£317	£1,808	-£1,491 (-£2,196, £840)
CVD events (MI, stroke, CHF)	£9,343	£9,812	-£469 (-£954, £44)
Adverse events	£248	£0	£248 (£79, £442)

Cost	SGLT2 inhibitors + standard care	Standard care	Incremental costs (95% CI)
Cost of death	£852	£1,105	-£253 (-£357, -£162)

I.6.2 Scenario analysis

Results of the scenario analysis are presented in Table I.13. SGLT2 inhibitors remain a cost-effective option in all the scenarios explored, with the exception of a scenario using a time horizon of 2.6 years (Scenario 2).

When evaluated over 2.6-year time horizon, the model predicted that very few patients have progressed to the more severe health states, and therefore the additional cost of SGLT2 inhibitors is not offset by the cost of averted transplants or from patients being in the more costly CKD stage health states and SGLT2 inhibitors are not cost-effective. However, when the analysis time frame was extended to a 5-year time horizon, the additional cost of SGLT2 inhibitors are almost offset, and there are sufficient cost savings that they become cost-effective.

SGLT2 inhibitors remain a dominant treatment option even when assuming a treatment waning effect at four years, which corresponds to the time at which the predicted mean eGFR of patients on SGLT2 inhibitors is approximately 45, which is when patients may discontinue treatment. In this case, there will be also a reduction in treatment acquisition costs which it was not possible to model. Since the model predicted cost savings in this scenario, a reduction in SGLT2 inhibitor acquisition costs would increase the cost savings.

Scenario analyses applying the CMU discount of canagliflozin were also conducted; however, the estimated total costs and ICER are confidential (reflect the inclusion of a confidential discount) and are not provided in Table I.13. In the base case analysis, SGLT2 inhibitors remained the dominant treatment option. When applied in the scenario where the time horizon was restricted to 2.6 years, the ICER remained over the threshold of cost-effectiveness (£20,000 per QALY).

Table I.13 Cost-effective results of the scenario analyses

	Total discounted costs	Total discounted QALYs	Incremental costs (95% CI)	Incremental QALYs (95% CI)	ICER (cost per QALY)
1. 40-year time horizon					
SGLT2 inhibitors + standard care	£72,243	6.23	-£3,759 (-£10,948, -£2,061)	0.95 (0.61, 1.29)	Intervention dominates
Standard care	£76,003	5.28	-	-	-
2. 2.6-year time horizon					

	Total discounted costs	Total discounted QALYs	Incremental costs (95% CI)	Incremental QALYs (95% CI)	ICER (cost per QALY)
SGLT2 inhibitors + standard care	£9,401	1.33	£888 (£440, £1,146)	0.01 (0.00, 0.01)	£166,992
Standard care	£813	1.33	-	-	-
3. 5-year time horizon					
SGLT2 inhibitors + standard care	£18,266	2.35	£237 (£575, £879)	0.06 (0.04, 0.07)	£4,275
Standard care	£18,029	2.30	-	-	-
4. Utilities for dialysis and transplant from Willis et al (2021)					
SGLT2 inhibitors + standard care	£35,467	3.75	-£3,709 (-£5,444, -£2,095)	0.30 (0.23, 0.36)	Intervention dominates
Standard care	£39,177	3.45	-	-	-
5. No costs and QoL impact for CKD stages					
SGLT2 inhibitors + standard care	£24,605	5.24	-£950 (-£2,857, £546)	0.25 (0.17, 0.34)	Intervention dominates
Standard care	£25,555	4.98	-	-	-
6. Treatment waning effect					
SGLT2 inhibitors + standard care	£37,018	3.80	-£1,916 (-£2,974, -£832)	0.14 (0.11, 0.18)	Intervention dominates
Standard care	£38,933	3.66	-	-	-
7. Alternative rates of transplantation					

	Total discounted costs	Total discounted QALYs	Incremental costs (95% CI)	Incremental QALYs (95% CI)	ICER (cost per QALY)
SGLT2 inhibitors + standard care	£35,678	3.89	-£3,702 (-£5,254, -£1,852)	0.24 (0.17, 0.3)	Intervention dominates
Standard care	£39,380	3.65	-	-	-

I.6.3 Subgroup analysis

The results of an exploratory analysis of patients with uACR 3-30 mg/mmol (population A2), presented in Table I.14, indicated that SGLT2 inhibitors may be cost-effective in this subgroup. In this analysis, SGLT2 inhibitors were both cost saving and associated with additional QALYs.

However, there were a number of limitations in the clinical evidence underpinning the economic analysis that meant that the results of this analysis are less reliable. Firstly, the progression of DKD, specifically eGFR over time, for patients on standard care was modelled using data for the population with uACR > 30 mg/mmol. The analysis of studies in the clinical evidence review suggested that there was a statistically significant difference in eGFR progression between uACR subgroups, therefore, disease progression may not be representative of patients in this subgroup. Secondly, there was less certainty in the clinical evidence of benefit in this subgroup, and the committee noted that not all studies reported outcomes for uACR populations in a consistent manner. Further to this, a recommendation in this subgroup may have a large potential resource impact due to the population size. The committee referred to the principle outlined in 6.2.14 of NICE's guide to the methods of technology appraisal (2013) and agreed that it would want to be increasingly certain of the cost-effectiveness of a technology as the resource impact of adoption increases.

Table I.14 Cost-effectiveness results of the subgroup analysis (A2 population)

	Total discounted costs	Total discounted QALYs	Incremental costs (95% CI)	Incremental QALYs (95% CI)	ICER (cost per QALY)
SGLT2 inhibitors + standard care	£28,594	4.06	-£3,578 (-£5,068, -£2,130)	0.21 (0.10, 0.31)	Intervention dominates
Standard care	£32,172	3.85	-	-	-

I.7 Discussion

The economic analysis estimated SGLT2 inhibitors likely to be cost-effective for patients with uACR > 30 mg/mmol, and found the results robust to a wide range of assumptions.

There were a number of limitations associated with the economic analysis, which generally stemmed from being constrained by the existing model structure that was being adapted for this guideline. Firstly, it was not possible to include the results of the evidence review on eGFR or uACR progression as these outcomes was not reported by the studies in the same way that was used in the economic analysis. It was also not possible to fully explore scenarios such as treatment discontinuation, which may occur when patients' eGFR reaches 45 mL/min/1.73m^2, as per the marketing authorisation for some of the SGLT2 inhibitors. However, the cost-effectiveness of SGLT2 inhibitors demonstrated across a wide range of scenarios provided reassurance that the results were robust despite these limitations.

The committee felt that there was uncertainty in the extrapolation of clinical outcomes beyond the follow-up period of the trials, specifically regarding the modelled predictions of eGFR decline beyond the duration of the trial from which the risk equations were estimated. This was because these were not supported by clinical evidence and were not validated, either against external evidence such as registry data or by the clinical experts working with the original model developers. From inspection of the extrapolated eGFR plots generated by the model, the committee also considered that the predictions may not be plausible, since eGFR progression is not thought to be linear. However, they understood that it was important to capture the downstream benefits of the intervention which occur after the duration of the trial using good-quality representative sources of external data and assumptions that are clinically valid. The downstream benefits had a large impact on the results of the analysis, and their exclusion from the analysis meant that the SGLT2 inhibitors do not appear cost-effective. A sensitivity analysis that limited the relative benefit of SGLT2 inhibitors after four years found that it remained a cost-effective use of resources. Assessment of the cost-effectiveness over different time horizons also supported the conclusions, unless a time horizon was used that was insufficiently long enough to capture the downstream benefits of SGLT2 inhibitors.

Limitations in the clinical evidence informing the subgroup analysis of patients with uACR 3-30 mg/mmol meant that the results of this exploratory analysis were less robust than that of the base-case analysis. These limitations included progression of DKD for patients on standard care modelled using data for the population with uACR > 30 mg/mmol, and less certainty in the clinical evidence of benefit. It was also not possible to update any of the other patient characteristics such as eGFR and age, which are likely to be correlated with uACR. A recommendation in this subgroup may have a large potential resource impact due to the population size, and it is necessary to be increasingly certain of the cost-effectiveness of a technology as the resource impact of adoption increases.

I.8 Conclusions

The economic analysis indicated that SGLT2 inhibitors are likely to be cost-effective (i.e. were associated with cost savings and additional QALYs, or result in an incremental cost-effectiveness ratio of less than £20,000 per QALY gained) in an analysis based on evidence for patients with uACR > 30 mg/mmol.

An exploratory analysis of patients with uACR 3-30 mg/mmol indicated a possibility for SGLT2 inhibitors to be cost-effective in this subgroup; however, the economic analysis was based on less robust evidence and a firm conclusion could not be made.

I.9 References

Alva ML, Gray A, Mihaylova B, Leal J, Holman RR. The impact of diabetes-related complications on healthcare costs: new results from the UKPDS (UKPDS 84). Diabetic Med. 2015;32(4):459–66.

Clarke P, Gray A, Holman R. Estimating utility values for health states of type 2 diabetic patients using the EQ-5D (UKPDS 62). Med Decis Making. 2002;22(4):340–9.

Curtis L, Burns A. (2019) Unit cost of health and social care 2019. University of Kent, UK

Jesky MD, Dutton M, Dasgupta I, et al. Health-related quality of life impacts mortality but not progression to end-stage renal disease in pre-dialysis chronic kidney disease: a prospective observational study. PLoS One. 2016;11(11):e0165675.

Lee AJ, Morgan CL, Conway P, Currie CJ. Characterisation and comparison of health-related quality of life for patients with renal failure. Curr Med Res Opin. 2005;21(11):1777–83.

Mahaffey KW, Jardine MJ, Bompoint S, et al. Canagliflozin and Cardiovascular and Renal Outcomes in Type 2 Diabetes Mellitus and Chronic Kidney Disease in Primary and Secondary Cardiovascular Prevention Groups. Circulation. 2019 Aug 27;140(9):739-750.

National Institute for Health and Care Excellence. Canagliflozin, dapagliflozin and empagliflozin monotherapy for treating type 2 diabetes— Assessment Report National Institute for Health and Care Excellence. 2015.

National Institute for Health and Care Excellence. Chronic kidney disease: assessment and management (update). Expected publication: 25 August 2021.

National Institute for Health and Care Excellence. Cinacalcet for the treatment of secondary hyperparathyroidism in patients with end-stage renal disease on maintenance dialysis therapy (TA117). Manufacturer Submission. 2007.

National Institute for Health and Care Excellence. Tolvaptan for treating autosomal dominant polycystic kidney disease (TA358) Manufacturer Submission. 2012.

NHS Digital (2019) Prescription Cost Analysis - England, 2018. Accessed at: https://digital.nhs.uk/data-and-information/publications/statistical/prescription-cost-analysis/2018

Peasgood T, Brennan A, Mansell P, Elliott J, Basarir H, Kruger J. The impact of diabetes-related complications on preference-based measures of health related quality of life in adults with type I diabetes. Med Decis Making. 2016;36(8):1020–33.

Perkovic V, Jardine MJ, Neal B, et al. Canagliflozin and renal outcomes in type 2 diabetes and nephropathy. N Engl J Med. 2019;380(24): 2295–306.

Shingler S, Fordham B, Evans M, et al. Utilities for treatment-related adverse events in type 2 diabetes. J Med Econ. 2015;18(1):45–55.

Sullivan PW, Ghushchyan VH. EQ-5D scores for diabetes-related comorbidities. Value Health. 2016;19(8):1002–8.

UK Renal Registry. 22nd Annual Report—data to 31/12/2018, Bristol, UK 2020. https://renal.org/audit-research/annual-report/22nd-annual-report-data-31122018. Accessed 12 Jul 2021.

Willis M, Nilsson A, Kellerborg K, Ball P. Cost-Effectiveness of Canagliflozin Added to Standard of Care for Treating Diabetic Kidney Disease (DKD) in Patients with Type 2 Diabetes Mellitus (T2DM) in England: Estimates Using the CREDEM-DKD Model. Diabetes Ther. 2021;12:313–328.

Willis M, Asseburg C, Slee A, Nilsson A, Neslusan C. Development and internal validation of a discrete event simulation model of diabetic kidney disease using CREDENCE trial data. Diabetes Ther. 2020;11(11):2657–76.

Appendix J – Excluded studies

Study	Reason
Anonymous. (2015) Correction to Efficacy And Safety Of Empagliflozin Added To Existing Antidiabetes Treatment In Patients With Type 2 Diabetes And Chronic Kidney Disease: A Randomised, Double-Blind, Placebo-Controlled Trial [Lancet Diabetes Endocrinol] 2014; 2: 369-84. The Lancet Diabetes and Endocrinology 3(3): e2	- Not a peer-reviewed publication
Anonymous. (2019) Corrigendum to: Efficacy and safety of dapagliflozin in patients with type 2 diabetes and moderate renal impairment (chronic kidney disease stage 3A): The DERIVE Study (Diabetes, Obesity and Metabolism, (2018), 20, 11, (2532-2540), 10.1111/dom.13413). Diabetes, Obesity and Metabolism 21(1): 203	- Not a peer-reviewed publication
Barkas, Fotios, Ntekouan, Sebastian Filippas, Liberopoulos, Evangelos et al. (2021) Sodium-Glucose Cotransporter-2 Inhibitors and Protection Against stroke in Patients with type 2 Diabetes and Impaired Renal Function: A Systematic Review and Meta-Analysis. Journal of Stroke and Cerebrovascular Diseases 30(5): 105708	- Wrong population *Not restricted to people with CKD*
Barnett, Anthony H, Mithal, Ambrish, Manassie, Jenny et al. (2014) Efficacy and safety of empagliflozin added to existing antidiabetes treatment in patients with type 2 diabetes and chronic kidney disease: a randomised, double-blind, placebo-controlled trial. The lancet. Diabetes & endocrinology 2(5): 369-84	- Study does not contain a relevant intervention *Dosage not BNF indicated*
Bhatia, Kirtipal, Fox, Arieh, Jain, Vardhmaan et al. (2021) Prevention of heart failure events with sodium-glucose co-transporter 2 inhibitors across a spectrum of cardio-renal-metabolic risk. European Journal of Heart Failure	- Wrong population *No limited to people with CKD.*
Butler, Javed, Zannad, Faiez, Fitchett, David et al. (2019) Empagliflozin Improves Kidney Outcomes in Patients With or Without Heart Failure. Circulation. Heart failure 12(6): e005875	- Wrong population *Does not report data for CKD subgroup.*
Cannon, Christopher P, Perkovic, Vlado, Agarwal, Rajiv et al. (2020) Evaluating the Effects of Canagliflozin on Cardiovascular and	- Secondary publication of an included study that does not provide any additional relevant information

Study	Reason
Renal Events in Patients With Type 2 Diabetes Mellitus and Chronic Kidney Disease According to Baseline HbA1c, Including Those With HbA1c <7%: Results From the CREDENCE Trial. Circulation 141(5): 407-410	*Reports outcomes subgrouped by baseline HBA1c*
Cherney, D (2016) The effect of sodium glucose cotransporter 2 inhibition with empagliflozin on microalbuminuria and macroalbuminuria in patients with type 2 diabetes. Diabetologia. 59 (9) (pp 1860-1870), 2016. Date of publication: 01 sep 2016.	- Pooled analysis of EMPA-REG data already included
Dekkers, Claire C. J., Wheeler, David C, David Sjostrom, C. et al. (2018) Erratum: Effects of the sodium-glucose co-transporter 2 inhibitor dapagliflozin in patients with type 2 diabetes and Stages 3b-4 chronic kidney disease (Nephrology Dialysis Transplantation (gfx350) DOI: 10.1093/NDT/gfx350). Nephrology Dialysis Transplantation 33(7): 1280	- Not a peer-reviewed publication
Ferreira, Joao Pedro, Zannad, Faiez, George, Jyothis T. et al. (2021) Cardio/Kidney Composite End Points: A Post Hoc Analysis of the EMPA-REG OUTCOME Trial. Journal of the American Heart Association: e020053	- Wrong population *Does not stratify by CKD status.*
Halalau, Alexandra; Fuller, William; Wheeler, Stephanie (2021) Canagliflozin Reduces the Risk of Kidney Failure in Patients with Type 2 Diabetes Mellitus and Nephropathy: the CREDENCE Randomized Trial. Journal of General Internal Medicine	- Not a peer-reviewed publication *Commentary*
Haneda, M, Seino, Y, Inagaki, N et al. (2016) Influence of Renal Function on the 52-Week Efficacy and Safety of the Sodium Glucose Cotransporter 2 Inhibitor Luseogliflozin in Japanese Patients with Type 2 Diabetes Mellitus. Clinical therapeutics 38(1): 66-88.e20	- Study does not contain a relevant intervention *Luseogliflozin does not have a UK marketing authorisation*
Heerspink, Hiddo J L, Stefansson, Bergur V, Correa-Rotter, Ricardo et al. (2020) Dapagliflozin in Patients with Chronic Kidney Disease. The New England journal of medicine 383(15): 1436-1446	- Wrong population *Population not restricted to type 2 diabetes and does not report type 2 subgroup (see Wheeler 2021 for type 2 subgroup)*
Ingelfinger, Julie R. and Rosen, Clifford J. (2019) Clinical credence - SGLT2 inhibitors, diabetes, and chronic kidney disease. New	- Not a peer-reviewed publication *Commentary*

Study	Reason
England Journal of Medicine 380(24): 2371-2373	
Jhund, Pardeep S., Docherty, Kieran F., McMurray, John J V et al. (2021) Efficacy of Dapagliflozin on Renal Function and Outcomes in Patients With Heart Failure With Reduced Ejection Fraction: Results of DAPA-HF. Circulation 143(4): 298-309	- Wrong population
Kohan, DE, Fioretto, P, Johnsson, K et al. (2016) The effect of dapagliflozin on renal function in patients with type 2 diabetes. Journal of nephrology 29(3): 391-400	- Wrong population *Did not include people with CKD.*
Kohan, Donald E., Fioretto, Paola, Tang, Weihua et al. (2014) Long-term study of patients with type 2 diabetes and moderate renal impairment shows that dapagliflozin reduces weight and blood pressure but does not improve glycemic control. Kidney International 85(4): 962-971	- Data not reported in an extractable format *Relevant outcomes not reported for subgroup CKD (eGFR <60) population*
Kraus, BJ, Weir, MR, Bakris, GL et al. (2020) Characterization and implications of the initial estimated glomerular filtration rate 'dip' upon sodium-glucose co-transporter-2 inhibition with empagliflozin in the EMPA-REG OUTCOME trial. Kidney international	- Secondary publication of an included study that does not provide any additional relevant information
Mahaffey, Kenneth W., Bompoint, Severine, Neal, Bruce et al. (2020) Canagliflozin and Cardiovascular and Renal Outcomes in Type 2 Diabetes Mellitus and Chronic Kidney Disease in Primary and Secondary Cardiovascular Prevention Groups: Results from the Randomized CREDENCE Trial. Circulation: 739-750	- Duplicate reference
Mayer, Gert J, Wanner, Christoph, Weir, Matthew R et al. (2019) Analysis from the EMPA-REG OUTCOME R trial indicates empagliflozin may assist in preventing the progression of chronic kidney disease in patients with type 2 diabetes irrespective of medications that alter intrarenal hemodynamics. Kidney international 96(2): 489-504	- Wrong population *Includes people with and without CKD and does not stratify results.*
Mosenzon, Ofri, Wiviott, Stephen D, Cahn, Avivit et al. (2019) Effects of dapagliflozin on development and progression of kidney disease	- Wrong population *Participants were required to have a eGFR>60*

Study	Reason
in patients with type 2 diabetes: an analysis from the DECLARE-TIMI 58 randomised trial. The lancet. Diabetes & endocrinology 7(8): 606-617	
Nieto Iglesias, J, Yale, JF, Bakris, G et al. (2013) Efficacy and safety of canagliflozin in subjects with type 2 diabetes mellitus and chronic kidney disease over 52 weeks. Diabetologia 56(suppl1): 381	- Conference abstract
Ohkuma, T, Van Gaal, L, Shaw, W et al. (2019) Clinical outcomes with canagliflozin according to baseline body mass index: results from post hoc analyses of the CANVAS Program. Diabetes, obesity & metabolism	- Wrong population *Does not stratify results for people with and without CKD.*
Okunrintemi, Victor, Mishriky, Basem M., Powell, James R. et al. (2021) Sodium-glucose co-transporter-2 inhibitors and atrial fibrillation in the cardiovascular and renal outcome trials. Diabetes, Obesity and Metabolism 23(1): 276-280	- Wrong population
Oshima, Megumi, Neal, Bruce, Toyama, Tadashi et al. (2020) Different eGFR Decline Thresholds and Renal Effects of Canagliflozin: Data from the CANVAS Program. Journal of the American Society of Nephrology : JASN 31(10): 2446-2456	- Wrong population *Population was not limited to people with CKD and no CKD subgroup reported.*
Oshima, Megumi, Neuen, Brendon L, Jardine, Meg J et al. (2020) Effects of canagliflozin on anaemia in patients with type 2 diabetes and chronic kidney disease: a post-hoc analysis from the CREDENCE trial. The lancet. Diabetes & endocrinology 8(11): 903-914	- Outcome - not within protocol *Reports outcomes related to anaemia*
Palmer, Suetonia C, Tendal, Britta, Mustafa, Reem A et al. (2021) Sodium-glucose cotransporter protein-2 (SGLT-2) inhibitors and glucagon-like peptide-1 (GLP-1) receptor agonists for type 2 diabetes: systematic review and network meta-analysis of randomised controlled trials. BMJ (Clinical research ed.) 372: m4573	- Wrong population
Patoulias, Dimitrios, Papadopoulos, Christodoulos, Stavropoulos, Konstantinos et al. (2021) Meta-analysis of Dedicated Renal Outcome Trials Assessing the Cardio-renal	- Review article but not a systematic review

Study	Reason
Efficacy of Sodium-Glucose Co-transporter-2 Inhibitors in Patients With Chronic Kidney Disease and Albuminuria. The American journal of cardiology 138: 116-118	
Perkovic, Vlado, Koitka-Weber, Audrey, Cooper, Mark E et al. (2020) Choice of endpoint in kidney outcome trials: considerations from the EMPA-REG OUTCOME R trial. Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association 35(12): 2103-2111	- Secondary publication of an included study that does not provide any additional relevant information
Perkovic, Vlado, Pfarr, Egon, Woerle, Hans J. et al. (2021) Choice of endpoint in kidney outcome trials: Considerations from the EMPA-REG OUTCOMEVR trial. Nephrology Dialysis Transplantation 35(12): 2103-2111	- Duplicate reference
Qu, Wei, Yao, Li, Liu, Xiaodan et al. (2021) Effects of Sodium-Glucose Co-transporter 2 Inhibitors on Hemoglobin Levels: A Meta-analysis of Randomized Controlled Trials. Frontiers in Pharmacology 12: 630820	- Outcome - not within protocol
Raji, Annaswamy, Xu, Zhi Jin, Lam, Raymond L. H. et al. (2020) Efficacy and Safety of Sitagliptin Compared with Dapagliflozin in People >= 65 Years Old with Type 2 Diabetes and Mild Renal Insufficiency. Diabetes Therapy	- Wrong population *Mean eGFR >60*
Salah, Husam M., Al'Aref, Subhi J., Al-Hawwas, Malek et al. (2021) Effect of sodium-glucose cotransporter 2 inhibitors on cardiovascular and kidney outcomes-Systematic review and meta-analysis of randomized placebo-controlled trials: SGLT2i-Cardiovascular and Kidney Outcomes. American Heart Journal 232: 10-22	- Duplicate reference
Salah, Husam M, Al'Aref, Subhi J, Khan, Muhammad Shahzeb et al. (2021) Effect of sodium-glucose cotransporter 2 inhibitors on cardiovascular and kidney outcomes-Systematic review and meta-analysis of randomized placebo-controlled trials. American heart journal 232: 10-22	- Wrong population *Analysis not presented for people required to have both T2D and CKD*
Sarraju, Ashish, Li, JingWei, Cannon, Christopher P et al. (2021) Effects of canagliflozin on cardiovascular, renal, and	- Secondary publication of an included study that does not provide any additional relevant information

Study	Reason
safety outcomes in participants with type 2 diabetes and chronic kidney disease according to history of heart failure: Results from the CREDENCE trial. American heart journal 233: 141-148	*Subgroup analysis with and without heart failure.*
Scott, Russell, Morgan, Jerry, Zimmer, Zachary et al. (2018) A randomized clinical trial of the efficacy and safety of sitagliptin compared with dapagliflozin in patients with type 2 diabetes mellitus and mild renal insufficiency: The CompoSIT-R study. Diabetes, obesity & metabolism 20(12): 2876-2884	- Wrong population *Population was people with mild renal insufficiency but not CKD*
Stefansson, Bergur V., Sjostrom, C. David, Heerspink, Hiddo J.L. et al. (2020) Correction of anemia by dapagliflozin in patients with type 2 diabetes. Journal of Diabetes and its Complications 34(12): 107729	- Wrong population *Mixed with and without CKD population - no subgroup analysis on those with.*
Takashima, Hiroyuki, Yoshida, Yoshinori, Nagura, Chinami et al. (2018) Renoprotective effects of canagliflozin, a sodium glucose cotransporter 2 inhibitor, in type 2 diabetes patients with chronic kidney disease: A randomized open-label prospective trial. Diabetes & vascular disease research 15(5): 469-472	- Wrong time point *Data not reported at agreed time point*
Wanner, Christoph, Heerspink, Hiddo J L, Zinman, Bernard et al. (2018) Empagliflozin and Kidney Function Decline in Patients with Type 2 Diabetes: A Slope Analysis from the EMPA-REG OUTCOME Trial. Journal of the American Society of Nephrology : JASN 29(11): 2755-2769	- Outcome - not within protocol *Only reports eGFR slope for CKD subgroup*
Wheeler, David C, Stefansson, Bergur V, Batiushin, Mikhail et al. (2020) The dapagliflozin and prevention of adverse outcomes in chronic kidney disease (DAPA-CKD) trial: baseline characteristics. Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association 35(10): 1700-1711	- Outcome - not within protocol *See Wheeler 2021 for outcomes for DAPA-CKD in type 2 diabetes subgroup*
Xu, Lubin, Li, Yang, Lang, Jiaxin et al. (2017) Effects of sodium-glucose co-transporter 2 (SGLT2) inhibition on renal function and albuminuria in patients with type 2 diabetes: A systematic review and meta-analysis. PeerJ 2017(6): e3405	- Wrong population

Study	Reason
Yoshihara, Fumiki, Imazu, Miki, Hamasaki, Toshimitsu et al. (2018) An Exploratory Study of Dapagliflozin for the Attenuation of Albuminuria in Patients with Heart Failure and Type 2 Diabetes Mellitus (DAPPER). Cardiovascular drugs and therapy 32(2): 183-190	- Study protcol
Yu, Jie, Li, Jingwei, Leaver, Phillip J et al. (2021) Effects of canagliflozin on myocardial infarction: a post hoc analysis of the CANVAS Program and CREDENCE trial. Cardiovascular research	- Secondary publication of an included study that does not provide any additional relevant information

Appendix K – Research recommendations – full details

K.1.1 Research recommendations

1. What is the clinical and cost effectiveness of SGLT2 inhibitors in in adults with type 2 diabetes and chronic kidney disease, stratified across different ethnic groups?

2. What is the clinical and cost effectiveness of SGLT2 inhibitors in adults with type 2 diabetes, chronic kidney disease and a urine ACR of less than 3 mg/mmol?

K.1.2 Why this is important

K.1.2.1 Research recommendation 1

Some ethnic groups more likely to be at risk of macrovascular or microvascular complications for a given level of kidney function. The kidney function criteria at which using an SGLT2 inhibitor becomes clinically and cost effective may therefore differ by ethnicity. Inclusion of participants from different ethnic groups and stratification of randomised controlled trials by ethnicity is important to assess this.

K.1.2.2 Research recommendation 2

Some people with type 2 diabetes, chronic kidney disease and an ACR of less than 3mg/mmol may benefit from being given an SGLT2 inhibitor but evidence was not available for this population, so a recommendation was not made.

K.1.3 Rationale for research recommendation

K.1.3.1 Research recommendation 1

Importance to 'patients' or the population	This is important that people of different ethnicities with CKD and type 2 diabetes receive the same standard of care as the broader population.
Relevance to NICE guidance	Medium: the research is relevant to the recommendations in the guidance, but the research recommendations are not essential to future updates. SGLT2 inhibitors are considered in this guideline as a treatment option, but there is not enough data comparing benefits in people of different ethnicities.
Relevance to the NHS	The outcome could affect the types of treatment prescribed to these groups. The results of the research could be used to provide care that was better tailored to particular ethnic groups.
National priorities	High
Current evidence base	No data was identified in the evidence review that compared important outcomes across difference ethnicities.
Equality considerations	Consideration would have to be given to a broad enough range of ethnicities as reflected in the UK population.

K.1.3.2 Research recommendation 2

Importance to 'patients' or the population	People with CKD and an ACR lower than 3mg/mmol might benefit from SGLT2 inhibitors, but evidence was not available for this population group specifically.
Relevance to NICE guidance	Medium: the research is relevant to the recommendations in the guidance, but the research recommendations are not essential to future updates. SGLT2 inhibitors are considered in this guideline as a treatment option, but there is not enough data to recommend these for people with an ACR lower than 3mg/mmol.
Relevance to the NHS	The outcome could affect the types of treatment prescribed to this population. The number of people with A1 CKD is large and any recommendation for SGLT2 inhibitors in this group would have a large resource impact. It is therefore particularly important to have good evidence on effectiveness and cost effectiveness to inform recommendations.
National priorities	High
Current evidence base	No data was identified in the evidence review that looked at the effectiveness of SGLT2 inhibitors in people with an ACR lower than 3mg/mmol with CKD.
Equality considerations	Consideration would have to be given to all groups with protected characteristics as outlined in the equality impact assessment published with this guideline

K.1.4 Modified PICO table

K.1.4.1 Research recommendation 1

Population	Adults with type 2 diabetes, chronic kidney disease from a range of ethnic groups that are representative of the UK population.
Intervention	SGLT2 inhibitors in addition to standard care
Comparator	Placebo in addition to standard care
Outcome	Cardiovascular events, Chronic kidney disease progression and adverse effect outcomes, including diabetic ketoacidosis, acute kidney injury and genitourinary infections
Study design	RCT. Stratification by ethnic group may be possible using individual patient data from existing randomised controlled trials.
Timeframe	>2 years

Additional information	None

K.1.4.2 Research recommendation 2

Population	Adults with type 2 diabetes, chronic kidney disease and an ACR of less than 3mg/mmol
Intervention	SGLT2 inhibitors in addition to standard care
Comparator	Placebo in addition to standard care
Outcome	Cardiovascular events, chronic kidney disease progression and adverse effect outcomes including diabetic ketoacidosis, acute kidney injury and genitourinary infections
Study design	RCT
Timeframe	>2years
Additional information	None

Appendix L – Methods

Review protocols

Review protocols were developed with the guideline committee to outline the inclusion and exclusion criteria used to select studies for each evidence review. Where possible, review protocols were prospectively registered in the PROSPERO register of systematic reviews.

Searching for evidence

Evidence was searched for each review question using the methods specified in the 2018 NICE guidelines manual.

Selecting studies for inclusion

All references identified by the literature searches and from other sources (for example, previous versions of the guideline or studies identified by committee members) were uploaded into EPPI reviewer software (version 5) and de-duplicated. Titles and abstracts were assessed for possible inclusion using the criteria specified in the review protocol. 10% of the abstracts were reviewed by two reviewers, with any disagreements resolved by discussion or, if necessary, a third independent reviewer.

As an additional check to ensure this approach did not miss relevant studies, systematic reviews (or qualitative evidence syntheses in the case of reviews of qualitative studies) were included in the review protocol and search strategy for all review questions. Relevant systematic reviews or qualitative evidence syntheses were used to identify any papers not found through the primary search. Committee members were also consulted to identify studies that were missed. If additional studies were found that were erroneously excluded during the priority screening process, the full database was subsequently screened.

The full text of potentially eligible studies was retrieved and assessed according to the criteria specified in the review protocol. A standardised form was used to extract data from included studies. Study investigators were contacted for missing data when time and resources allowed (when this occurred, this was noted in the evidence review and relevant data was included).

Incorporating published evidence syntheses

For all review questions where a literature search was undertaken looking for a particular study design, published evidence syntheses (quantitative systematic reviews or qualitative evidence syntheses) containing studies of that design were also included. All included studies from those syntheses were screened to identify any additional relevant primary studies not found as part of the initial search. Evidence syntheses that were used solely as a source of primary studies were not formally included in the evidence review (as they did not provide additional data) and were not quality assessed.

If published evidence syntheses were identified sufficiently early in the review process (for example, from the surveillance review or early in the database search), they were considered for use as the primary source of data, rather than extracting information from primary studies. Syntheses considered for inclusion in this way were quality assessed to assess their suitability using the appropriate checklist, as outlined in

Table 3. Note that this quality assessment was solely used to assess the quality of the synthesis in order to decide whether it could be used as a source of data, as outlined in

Table 4, not the quality of evidence contained within it, which was assessed in the usual way as outlined in the section on ‘Appraising the quality of evidence’.

Table 3: Checklists for published evidence syntheses

Type of synthesis	Checklist for quality appraisal
Systematic review of quantitative evidence	ROBIS
Network meta-analysis	Modified version of the PRISMA NMA tool (see appendix K of ‘Developing NICE guidelines, the manual’)
Qualitative evidence synthesis	ENTREQ reporting standard for published evidence synthesis (https://bmcmedresmethodol.biomedcentral.com/articles/10.1186/1471-2288-12-181) is the generic reporting standard for QES, however specific reporting standards exist for meta-ethnography (eMERGe [https://emergeproject.org/]) and for realist synthesis (RAMESES II [https://www.ramesesproject.org/]). If these reporting standards are not appropriate to the QES then an adapted PRISMA framework is used (see Flemming K, Booth A, Hannes K, Cargo M, Noyes J. Cochrane Qualitative and Implementation Methods Group guidance series-paper 6: reporting guidelines for qualitative, implementation, and process evaluation evidence syntheses. Journal of Clinical Epidemiology 2018; 97: 79-85).
Individual patient data meta-analysis	Checklist based on Tierney, Jayne F., et al. "Individual participant data (IPD) meta-analyses of randomised controlled trials: guidance on their use." PLoS Med 12.7 (2015): e1001855.

Each published evidence synthesis was classified into one of the following three groups:

- High quality – It is unlikely that additional relevant and important data would be identified from primary studies compared to that reported in the review, and unlikely that any relevant and important studies have been missed by the review.
- Moderate quality – It is possible that additional relevant and important data would be identified from primary studies compared to that reported in the review, but unlikely that any relevant and important studies have been missed by the review.
- Low quality – It is possible that relevant and important studies have been missed by the review.

Each published evidence synthesis was also classified into one of three groups for its applicability as a source of data, based on how closely the review matches the specified review protocol in the guideline. Studies were rated as follows:

- Fully applicable – The identified review fully covers the review protocol in the guideline.
- Partially applicable – The identified review fully covers a discrete subsection of the review protocol in the guideline (for example, some of the factors in the protocol only).
- Not applicable – The identified review, despite including studies relevant to the review question, does not fully cover any discrete subsection of the review protocol in the guideline.

The way that a published evidence synthesis was used in the evidence review depended on its quality and applicability, as defined in **Table 4**. When published evidence syntheses were used as a source of primary data, data from these evidence syntheses were quality assessed and presented in GRADE/CERQual tables in the same way as if data had been extracted from primary studies. In questions where data was extracted from both systematic reviews and primary studies, these were checked to ensure none of the data had been double counted through this process.

Table 4: Criteria for using published evidence syntheses as a source of data

Quality	Applicability	Use of published evidence synthesis
High	Fully applicable	Data from the published evidence synthesis were used instead of undertaking a new literature search or data analysis. Searches were only done to cover the period of time since the search date of the review. If the review was considered up to date (following discussion with the guideline committee and NICE lead for quality assurance), no additional search was conducted.
High	Partially applicable	Data from the published evidence synthesis were used instead of undertaking a new literature search and data analysis for the relevant subsection of the protocol. For this section, searches were only done to cover the period of time since the search date of the review. If the review was considered up to date (following discussion with the guideline committee and NICE lead for quality assurance), no additional search was conducted. For other sections not covered by the evidence synthesis, searches were undertaken as normal.
Moderate	Fully applicable	Details of included studies were used instead of undertaking a new literature search. Full-text papers of included studies were still retrieved for the purposes of data analysis. Searches were only done to cover the period of time since the search date of the review.
Moderate	Partially applicable	Details of included studies were used instead of undertaking a new literature search for the relevant subsection of the protocol. For this section, searches were only done to cover the period of time since the search date of the review. For other sections not covered by the evidence synthesis, searches were undertaken as normal.

Methods of combining evidence

Evidence synthesis and pairwise meta-analysis

Where possible, meta-analyses were conducted to combine the results of quantitative studies for each outcome. Network meta-analyses was considered in situations where there were at least 3 treatment alternatives. When there were 2 treatment alternatives, pairwise meta-analysis was used to compare interventions. A pooled relative risk was calculated for dichotomous outcomes (using the Mantel–Haenszel method) reporting numbers of people having an event, and a pooled incidence rate ratio was calculated for dichotomous outcomes reporting total numbers of events. Both relative and absolute risks were presented, with absolute risks calculated by applying the relative risk to the risk in the comparator arm of the meta-analysis (calculated as the total number events in the comparator arms of studies in the meta-analysis divided by the total number of participants in the comparator arms of studies in the meta-analysis).

A pooled mean difference was calculated for continuous outcomes (using the inverse variance method) when the same scale was used to measure an outcome across different studies. For continuous outcomes analysed as mean differences, change from baseline values were used in the meta-analysis if they were accompanied by a measure of spread (for example standard deviation). Where change from baseline (accompanied by a measure of spread) were not reported, the corresponding values at the timepoint of interest were used. If only a subset of trials reported change from baseline data, final timepoint values were combined with change from baseline values to produce summary estimates of effect.

Random effects models were fitted when there was significant between-study heterogeneity in methodology, population, intervention or comparator was identified by the reviewer in advance of data analysis. This decision was made and recorded before any data analysis was undertaken. For all other syntheses, fixed- and random-effects models were fitted, with the presented analysis dependent on the degree of heterogeneity in the assembled evidence. Fixed-effects models were the preferred choice to report, but in situations where the assumption of a shared mean for fixed-effects model were clearly not met, even after appropriate pre-specified subgroup analyses were conducted, random-effects results are presented. Fixed-effects models were deemed to be inappropriate if there was significant statistical heterogeneity in the meta-analysis, defined as $I^2 \geq 50\%$. However, in cases where the results from individual pre-specified subgroup analyses were less heterogeneous (with $I^2 < 50\%$) the results from these subgroups were reported using fixed effects models. This may have led to situations where pooled results were reported from random-effects models and subgroup results were reported from fixed-effects models. Results were reported separately in the GRADE profiles only when there was evidence for a difference in effect across subgroups (test for subgroup differences, $p<0.05$).

Quality assessment

RCTs and quasi-randomised controlled trials were quality assessed using the Cochrane Risk of Bias Tool. Evidence on each outcome for each individual study was classified into one of the following groups:

- Low risk of bias – The true effect size for the study is likely to be close to the estimated effect size.
- Moderate risk of bias – There is a possibility the true effect size for the study is substantially different to the estimated effect size.
- High risk of bias – It is likely the true effect size for the study is substantially different to the estimated effect size.

Each individual study was also classified into one of three groups for directness, based on if there were concerns about the population, intervention, comparator and/or outcomes in the study and how directly these variables could address the specified review question. Studies were rated as follows:

- Direct – No important deviations from the protocol in population, intervention, comparator and/or outcomes.
- Partially indirect – Important deviations from the protocol in one of the following areas: population, intervention, comparator and/or outcomes.
- Indirect – Important deviations from the protocol in at least two of the following areas: population, intervention, comparator and/or outcomes.

Minimally important differences (MIDs) and clinical decision thresholds

The Core Outcome Measures in Effectiveness Trials (COMET) database was searched to identify published minimal clinically important difference thresholds relevant to this guideline that might aid the committee in identifying clinical decision thresholds for the purpose of GRADE. Identified MIDs were assessed to ensure they had been developed and validated in a methodologically rigorous way, and were applicable to the populations, interventions and outcomes specified in this guideline. In addition, the Guideline Committee were asked to prospectively specify any outcomes where they felt a consensus clinical decision threshold could be defined from their experience. In particular, any questions looking to evaluate noninferiority (that one treatment is not meaningfully worse than another) required a clinical decision threshold to be defined to act as a non-inferiority margin.

Clinical decision thresholds were used to assess imprecision using GRADE and aid interpretation of the size of effects for different outcomes. No published mininimally important differences were found for this evidence review and no specific consensus decision thresholds were made.

For continuous outcomes expressed as a mean difference where no other clinical decision threshold was available, a clinical decision threshold of 0.5 of the median standard deviations of the comparison group arms was used (Norman et al. 2003). For continuous outcomes expressed as a standardised mean difference where no other clinical decision threshold was available, a clinical decision threshold of 0.5 standard deviations was used. For SMDs that were back converted to one of the original scales to aid interpretation, rating of imprecision was carried out before back calculation. For relative risks, where no other clinical decision threshold was available, a default clinical decision threshold for dichotomous outcomes of 0.8 to 1.25 was used. Odds ratios were converted to risk ratios before presentation to the committee to aid interpretation.

GRADE for intervention studies analysed using pairwise analysis

GRADE was used to assess the quality of evidence for the outcomes specified in the review protocol. Data from randomised controlled trials, non-randomised controlled trials and cohort studies (which were quality assessed using the Cochrane risk of bias tool or ROBINS-I) were initially rated as high quality while data from other study types were initially rated as low quality. The quality of the evidence for each outcome was downgraded or not from this initial point, based on the criteria given in Table 5.

Table 5: Rationale for downgrading quality of evidence for intervention studies

GRADE criteria	Reasons for downgrading quality
Risk of bias	Not serious: If less than 33.3% of the weight in a meta-analysis came from studies at moderate or high risk of bias, the overall outcome was not downgraded. Serious: If greater than 33.3% of the weight in a meta-analysis came from studies at moderate or high risk of bias, the outcome was downgraded one level. Very serious: If greater than 33.3% of the weight in a meta-analysis came from studies at high risk of bias, the outcome was downgraded two levels. Extremely serious: If greater than 33.3% of the weight in a meta-analysis came from studies at critical risk of bias, the outcome was downgraded three levels
Indirectness	Not serious: If less than 33.3% of the weight in a meta-analysis came from partially indirect or indirect studies, the overall outcome was not downgraded. Serious: If greater than 33.3% of the weight in a meta-analysis came from partially indirect or indirect studies, the outcome was downgraded one level. Very serious: If greater than 33.3% of the weight in a meta-analysis came from indirect studies, the outcome was downgraded two levels.
Inconsistency	Concerns about inconsistency of effects across studies, occurring when there is unexplained variability in the treatment effect demonstrated across studies (heterogeneity), after appropriate pre-specified subgroup analyses have been conducted. This was assessed using the I^2 statistic. N/A: Inconsistency was marked as not applicable if data on the outcome was only available from one study. Not serious: If the I^2 was less than 33.3%, the outcome was not downgraded. Serious: If the I^2 was between 33.3% and 66.7%, the outcome was downgraded one level.

GRADE criteria	Reasons for downgrading quality
	Very serious: If the I^2 was greater than 66.7%, the outcome was downgraded two levels.
Imprecision	If an MID other than the line of no effect was defined for the outcome, the outcome was downgraded once if the 95% confidence interval for the effect size crossed one line of the MID, and twice if it crosses both lines of the MID. If the line of no effect was defined as an MID for the outcome, it was downgraded once if the 95% confidence interval for the effect size crossed the line of no effect (i.e. the outcome was not statistically significant), and twice if the sample size of the study was sufficiently small that it is not plausible any realistic effect size could have been detected. Outcomes meeting the criteria for downgrading above were not downgraded if the confidence interval was sufficiently narrow that the upper and lower bounds would correspond to clinically equivalent scenarios.
Publication bias	Where 10 or more studies were included as part of a single meta-analysis, a funnel plot was produced to graphically assess the potential for publication bias. When a funnel plot showed convincing evidence of publication bias, or the review team became aware of other evidence of publication bias (for example, evidence of unpublished trials where there was evidence that the effect estimate differed in published and unpublished data), the outcome was downgraded once. If no evidence of publication bias was found for any outcomes in a review (as was often the case), this domain was excluded from GRADE profiles to improve readability.

For outcomes that were originally assigned a quality rating of ‘low’ (when the data was from observational studies that were not appraised using the ROBINS-I checklist), the quality of evidence for each outcome was upgraded if any of the following three conditions were met and the risk of bias for the outcome was rated as ‘no serious’:

- Data from studies showed an effect size sufficiently large that it could not be explained by confounding alone.
- Data showed a dose-response gradient.
- Data where all plausible residual confounding was likely to increase our confidence in the effect estimate.