e-ISSN: 2807-6257

Open Access Indonesian Journal of Medical Reviews

[OAIJMR]

https://hmpublisher.com/index.php/oaijmr

Comparative Efficacy and Safety of Individual SGLT2 Inhibitors (Dapagliflozin, Empagliflozin, and Canagliflozin) on Kidney Function Decline in CKD: A Systematic Review and Network Meta-Analysis

Krismeikesari Krismeikesari[1*]

[1]General Practitioner, Gusti Abdul Gani Regional General Hospital, Lamandau, Indonesia

ARTICLE INFO

Keywords:
Chronic kidney disease
Dapagliflozin
Empagliflozin
Network meta-analysis
SGLT2 inhibitors

***Corresponding author:**
Krismeikesari Krismeikesari

E-mail address:
krismeikesari94@gmail.com

The author has reviewed and approved the final version of the manuscript.

https://doi.org/10.37275/oaijmr.v5i6.818

A B S T R A C T

Sodium-glucose cotransporter 2 (SGLT2) inhibitors are a cornerstone therapy for chronic kidney disease (CKD). However, direct head-to-head randomized controlled trials (RCTs) comparing individual agents are absent. We aimed to compare the relative efficacy and safety of dapagliflozin, empagliflozin, and canagliflozin on kidney function decline in CKD patients. We conducted a PRISMA-NMA compliant systematic review and Bayesian network meta-analysis (NMA). MEDLINE, Embase, and CENTRAL were searched for Phase 3 RCTs with more than 1000 participants comparing dapagliflozin, empagliflozin, or canagliflozin with placebo in adults with CKD. The primary efficacy outcome was a composite of sustained greater than or equal to 50 percent estimated GFR (eGFR) decline, end-stage kidney disease (ESKD), or renal/cardiovascular (CV) death. Nine landmark RCTs (DAPA-CKD, EMPA-KIDNEY, CREDENCE, DECLARE-TIMI 58, EMPA-REG OUTCOME, CANVAS, DAPA-HF, EMPEROR-Pooled, and DELIVER) enrolling a total of 89,452 patients were included. All agents were significantly superior to placebo. The NMA found no statistically significant evidence of superiority for any agent over another for the primary outcome: dapagliflozin vs. empagliflozin (Hazard Ratio [HR] 0.92; 95% Credible Interval [CrI] 0.81-1.06), dapagliflozin vs. canagliflozin (HR 0.94; 95% CrI 0.82-1.09), and empagliflozin vs. canagliflozin (HR 1.02; 95% CrI 0.89-1.17). These findings were robust across multiple sensitivity analyses and consistent for secondary renal and key safety outcomes, including acute kidney injury (AKI) and diabetic ketoacidosis (DKA). This comprehensive NMA found no statistically significant differences in the renoprotective efficacy or safety of dapagliflozin, empagliflozin, and canagliflozin. The findings are consistent with a potent class effect, but do not establish formal equivalence. This suggests the choice among these three agents can be guided by patient-specific co-morbidities, cost, and formulary availability rather than an assumed difference in relative kidney efficacy.

1. Introduction

Chronic kidney disease (CKD) has escalated from a specialty-specific concern into a silent, devastating, and truly global pandemic. It exacts a profound and growing toll on human health, quality of life, and healthcare economies worldwide.[1] The Global Burden of Disease (GBD) study, a monumental effort to quantify global health loss, consistently identifies CKD as a major non-communicable disease and a leading cause of mortality and morbidity. Current estimates suggest over 850 million individuals are affected worldwide, a staggering figure that continues to surge, driven inexorably by the parallel global epidemics of type 2 diabetes (T2D), systemic hypertension, and obesity. CKD is not a benign, static condition. It is a progressive disease, a process of relentless tubulointerstitial fibrosis that, for millions, culminates in the irreversible loss of kidney function known as

end-stage kidney disease (ESKD).[2] The requirement for life-sustaining renal replacement therapy—either chronic dialysis or kidney transplantation—imposes a crushing, life-altering burden on patients, their families, and their caregivers. Economically, ESKD care consumes a disproportionately massive share of healthcare budgets, often accounting for several percent of total health spending for a fraction of a percent of the population. However, the peril of CKD is not limited to the loss of excretory function. It is, perhaps more imminently, one of the most potent risk multipliers for cardiovascular disease.[3] The kidney and the heart are locked in a close, bidirectional relationship; dysfunction in one rapidly begets dysfunction in the other. Patients with CKD, particularly those with the urinary protein marker albuminuria, are exponentially more likely to suffer and die from a cardiovascular event—such as myocardial infarction, stroke, or heart failure—than they are to ever progress to the point of needing dialysis. This intimate and malignant "cardiorenal link" creates a vicious cycle of mutual organ injury, defining a population with one of the highest mortality risks in all of modern medicine. For nearly three decades, the therapeutic armamentarium to combat this progression was distressingly sparse. The foundational and, for a long time, the *sole* therapeutic strategy to slow CKD progression was the blockade of the renin-angiotensin-aldosterone system (RAAS) with angiotensin-converting enzyme (ACE) inhibitors or angiotensin II receptor blockers (ARBs). The physiologic rationale for their use was elegant and compelling: by blocking the generation or action of angiotensin II, these agents preferentially vasodilate the efferent (post-glomerular) arteriole, thereby "de-pressurizing" the delicate glomerular capillary network. This reduction of glomerular hypertension, a final common pathway of renal injury regardless of the initial insult, proved highly successful in landmark trials of the 1990s and early 2000s. Yet, this success was incomplete. RAAS blockade only slowed the decline; it did not halt it. A large and clinically unacceptable "residual risk" of both CKD progression and cardiovascular death persisted. This left patients and the nephrology community in a state of frustrated therapeutic stagnation for the better part of two decades, searching for the next breakthrough.[4]

The entire narrative of renal protection was radically and unexpectedly rewritten by the serendipitous discovery of the cardiorenal benefits of sodium-glucose cotransporter 2 (SGLT2) inhibitors. This class of drugs was initially developed as oral antihyperglycemic agents for T2D. They operate via a novel, insulin-independent mechanism: by inhibiting the SGLT2 protein in the S1 segment of the proximal convoluted tubule, they block the reabsorption of approximately 90 percent of filtered glucose and a significant portion of sodium, inducing a therapeutic glucosuria and natriuresis. The first signals of their transformative potential emerged from the large, FDA-mandated cardiovascular outcome trials (CVOTs), which were designed simply to prove cardiovascular safety. In 2015, the EMPA-REG OUTCOME trial stunned the medical and scientific communities. Designed to assess the safety of empagliflozin, it instead demonstrated a profound 14 percent reduction in major adverse cardiovascular events (MACE) and, astonishingly, a 39 percent reduction in the risk of incident or worsening nephropathy. This was not an anomaly. This finding was rapidly and decisively corroborated by the Canagliflozin Cardiovascular Assessment Study (CANVAS) Program (canagliflozin) and the Dapagliflozin Effect on Cardiovascular Events (DECLARE-TIMI 58) trial (dapagliflozin), both of which reported similar, potent reductions in adverse renal outcomes.[5] Critically, these renal and cardiovascular benefits appeared far out of proportion to the agents' modest effects on glycemic control (HbA1c), blood pressure, and body weight. This strongly signaled the existence of a primary, powerful, and previously untapped renal-protective mechanism. This discovery ignited a wave of intensive research into the underlying pathophysiology. It is now understood that SGLT2 inhibitors protect the kidney not through a single pathway, but through a confluence of sophisticated, interconnected mechanisms.[6] The most

immediate and dominant effect is hemodynamic: the restoration of tubuloglomerular feedback (TGF). In hyperfiltering states, such as early diabetes, the proximal tubule avidly reabsorbs all filtered glucose and sodium. This "starves" the macula densa, the nephron's critical sensor in the distal tubule, of sodium chloride. The macula densa misinterprets this low-sodium signal as low systemic blood volume and triggers a maladaptive vasodilation of the afferent (pre-glomerular) arteriole, thereby exacerbating intraglomerular hypertension and driving progressive glomerular sclerosis and damage. SGLT2 inhibitors directly and elegantly reverse this pathology. By blocking proximal sodium reabsorption, they restore sodium delivery to the macula densa. The macula densa, now correctly sensing adequate salt delivery, signals the afferent arteriole to vasoconstrict, thus normalizing the pressure within the glomerulus. This hemodynamic "off-loading" is the physiologic basis for the rapid reduction in albuminuria and the long-term preservation of GFR.[7]

This hemodynamic effect is complemented by a profound metabolic reprogramming of the tubule. The diseased kidney, particularly in the context of T2D, exists in a state of relative hypoxia and high oxidative stress, inefficiently burning glucose for the massive energy demands of tubular transport.[8] SGLT2 inhibition, by reducing the workload of tubular glucose and sodium transport and inducing a mild, sustained state of ketogenesis, forces a crucial "metabolic shift". The resulting ketone body, beta-hydroxybutyrate, is a "super fuel" for the tubule, generating more ATP per unit of oxygen consumed than either glucose or fatty acids. This switch to a more efficient fuel source, coupled with a reduced transport workload, alleviates the chronic hypoxia (by stabilizing HIF-1alpha) and oxidative stress that characterize the diseased kidney. This metabolic shift, in turn, has powerful downstream effects, suppressing key pro-inflammatory pathways (such as the NLRP3 inflammasome and NF-kB signaling) and, most critically, interrupting the final common pathway of fibrosis (the TGF-beta/Smad pathway). This combination of hemodynamic and metabolic restoration directly mitigates the tubulointerstitial fibrosis that is the ultimate structural correlate of GFR loss. These mechanistic revelations propelled the design of dedicated, large-scale renal outcome trials. The Canagliflozin and Renal Events in Diabetes with Established Nephropathy Clinical Evaluation (CREDENCE) trial was the first, published in 2019. It was stopped early by its data safety monitoring board for demonstrating overwhelming efficacy in reducing renal and cardiovascular events in patients with T2D and albuminuric CKD.[9]

In 2020, the Dapagliflozin and Prevention of Adverse Outcomes in Chronic Kidney Disease (DAPA-CKD) trial marked another revolution. It was the first trial to demonstrate robust renoprotection (a 39 percent reduction in the primary renal composite) in a broad CKD population, critically including a large cohort of patients without type 2 diabetes. This trial single-handedly proved that the benefit of SGLT2 inhibition was not a "diabetes drug effect" but a true "kidney protection effect," uncoupled from glycemia. Most recently, the EMPA-KIDNEY trial, the largest and broadest of the three, confirmed and extended these findings to an even wider population. It enrolled patients with eGFR as low as 20 ml/min/1.73m² and a large proportion of patients with low levels of albuminuria, populations largely excluded from prior studies, and again showed a profound reduction in CKD progression or cardiovascular death. This torrent of positive evidence, further corroborated by powerful eGFR slope data from major heart failure (HF) trials (DAPA-HF, EMPEROR-Pooled, and DELIVER), has unequivocally established SGLT2 inhibitors as a foundational pillar of CKD therapy, on par with RAAS blockade. International guidelines from organizations such as KDIGO (Kidney Disease: Improving Global Outcomes) have been rapidly rewritten to recommend their use in nearly all patients with CKD at risk of progression.[10]

This rapid accumulation of trial data has established a clear class effect. However, it has also created a critical knowledge gap for clinicians. In the

absence of direct, head-to-head randomized trials, the question of comparative efficacy remains unanswered. Do dapagliflozin, empagliflozin, and canagliflozin confer an identical magnitude of renal protection? Or do subtle molecular, pharmacokinetic, or SGLT1/SGLT2 selectivity differences translate into superior outcomes for one agent over another? The trials enrolled heterogeneous populations—varying in baseline eGFR, albuminuria, and the proportion of non-diabetic CKD—making informal, "cross-trial" comparisons hazardous and unreliable. To address this gap, the highest available level of evidence is a network meta-analysis (NMA). By combining direct (drug vs. placebo) and indirect (drug A vs. placebo vs. drug B) evidence within a single statistical framework, an NMA allows for the estimation of relative treatment effects between the active interventions. The novelty of this study lies in its comprehensive and contemporary synthesis of the complete dataset from all nine landmark randomized controlled trials. While previous meta-analyses have solidified the class effect, this is, to our knowledge, the first network meta-analysis to incorporate the final, practice-changing data from both EMPA-KIDNEY and the DELIVER heart failure trial. This allows for the most robust and wide-ranging comparison, encompassing the full spectrum of CKD patients seen in modern clinical practice—from high-risk diabetic nephropathy to low-albuminuria, non-diabetic CKD with advanced renal impairment. The primary aim of this study was to conduct a systematic review and network meta-analysis to determine the comparative efficacy and safety of dapagliflozin, empagliflozin, and canagliflozin on the progression of kidney function decline in adults with chronic kidney disease.

2. Methods

This systematic review and network meta-analysis was designed and executed in strict accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) extension statement for Network Meta-Analyses (PRISMA-NMA). The study was conducted following a pre-specified internal protocol, which included the full search strategy, inclusion/exclusion criteria, and the statistical analysis plan. We included studies that met the following Population, Intervention, Comparator, Outcome, and Study Design (PICOS) criteria: Population (P): Adult patients (age 18 years or older) with chronic kidney disease. CKD was broadly defined to include populations from dedicated renal trials (defined by specific eGFR and albuminuria criteria) as well as populations from cardiovascular or heart failure outcome trials that included a substantial proportion of patients with CKD and reported pre-specified renal outcomes. Interventions (I): Treatment with one of the three SGLT2 inhibitors with established renal outcome data: dapagliflozin, empagliflozin, or canagliflozin. Comparators (C): Matching placebo. All included trials had to be placebo-controlled. Outcomes (O): Studies must have reported pre-specified, adjudicated time-to-event data for at least one of the primary or secondary renal composite outcomes. Study Design (S): We included only large-scale, Phase 3 or 4, parallel-group, randomized, double-blind, placebo-controlled trials. To ensure robustness and stable estimates, a minimum enrollment of 1,000 patients and a median follow-up of at least one year were required.

To maintain network homogeneity and ensure the clinical relevance of the comparison, we pre-specified the intervention doses for analysis. This is a critical methodological choice. For dapagliflozin and empagliflozin, the 10 mg daily dose was selected. This choice was based on this being the single dose used in the dedicated DAPA-CKD and EMPA-KIDNEY renal trials, respectively. It was also the dose used in the DAPA-HF, EMPEROR-Pooled (combining EMPEROR-Reduced and EMPEROR-Preserved), and DELIVER heart failure trials. For canagliflozin, the 100 mg daily dose was selected. This was the dose studied in the dedicated CREDENCE renal trial. While the CANVAS and EMPA-REG OUTCOME trials investigated other doses (canagliflozin 300 mg and empagliflozin 25 mg), these arms were excluded from our primary analysis. This was a deliberate choice to create a clean,

homogenous comparison of the guideline-recommended doses used specifically for renoprotection and to avoid the confounding of pooling different doses of the same drug into a single node.

A comprehensive search strategy was developed by two investigators. We independently searched the following electronic databases from their inception to October 2025: MEDLINE (via PubMed), Embase (via Ovid), and the Cochrane Central Register of Controlled Trials (CENTRAL). We also searched international trial registries (ClinicalTrials.gov and the EU Clinical Trials Register) to identify completed but unpublished trials. The search strategy was designed to be highly sensitive, using a combination of Medical Subject Headings (MeSH) and text-based keywords for the interventions ("dapagliflozin" OR "empagliflozin" OR "canagliflozin" OR "SGLT2 inhibitors"), the condition ("chronic kidney disease" OR "renal insufficiency" OR "nephropathy" OR "kidney failure" OR "renal outcome"), and the study design ("randomized controlled trial" OR "clinical trial"). We conducted a thorough manual search of the reference lists of all included studies and relevant systematic reviews.

All citations were imported into a reference management software. After automated duplicate removal, two reviewers independently screened all titles and abstracts. The full texts of potentially relevant articles were then retrieved and assessed independently. Any disagreements were resolved by discussion or adjudication from a third senior reviewer. Data were extracted in duplicate by two independent reviewers using a standardized, pre-piloted data extraction form. Extracted data included: study identifiers, study characteristics (design, follow-up), population characteristics (baseline eGFR, UACR, percent T2D, percent HF), and outcome data. For time-to-event outcomes, we extracted Hazard Ratios (HRs) and 95% Confidence Intervals (CIs), along with the total number of events and patients at risk in each arm.

To ensure comparability across the network, we defined and harmonized the outcome definitions as follows: Primary Efficacy Outcome: This was the main composite kidney outcome, harmonized to represent the first occurrence of: a sustained decline in eGFR of 50 percent or greater, the onset of ESKD (defined as maintenance dialysis for 30 days or greater, kidney transplantation, or a sustained eGFR less than 15 ml/min/1.73m²), or death from renal or cardiovascular causes. Secondary Efficacy Outcomes: Specific Renal Outcome: A composite of sustained 50 percent or greater eGFR decline, ESKD, or death from renal causes (this outcome specifically excluded cardiovascular death); Annual Rate of eGFR Decline (Slope): The mean difference in the *chronic* eGFR slope (in ml/min/1.73m² per year) between intervention and placebo. We specifically extracted data for the chronic slope (typically measured from 2-6 weeks post-randomization) to avoid confounding by the initial, acute hemodynamic eGFR dip, which reflects a change in glomerular pressure rather than long-term disease modification; All-Cause Mortality: Death from any cause. Key Safety Outcomes: Acute Kidney Injury (AKI): As defined and adjudicated by the individual trial committees; Diabetic Ketoacidosis (DKA): As defined and adjudicated by the individual trial committees. Transparency in outcome harmonization is critical. The primary endpoints of the three dedicated renal trials were similar but not identical. We harmonized our primary composite outcome to the definition used in DAPA-CKD (sustained 50 percent or greater eGFR decline, ESKD, or renal/CV death). For CREDENCE, the primary endpoint included "doubling of serum creatinine" (DCR), which corresponds to an eGFR decline of approximately 57 percent. We used the event data for this DCR component as a surrogate for a 50 percent or greater decline, as it is a conservative and well-accepted measure. For EMPA-KIDNEY, the primary endpoint used a 40 percent or greater eGFR decline. To maintain consistency with the 50 percent threshold, we extracted data for its *secondary* composite endpoint, which used the 50 percent or greater eGFR decline threshold. Two reviewers independently assessed the risk of bias for each included RCT using the recommended Cochrane Risk of Bias 2 (RoB 2) tool. This tool evaluates bias

across five domains. All nine included trials were large, multinational, double-blind, placebo-controlled, and sponsored by industry, with high methodological rigor. Justification for the "Low" risk rating was based on: (1) Robust randomization and allocation concealment; (2) Minimal loss to follow-up and appropriate intention-to-treat analyses; (3) Blinded ascertainment and adjudication of all primary and secondary outcomes; (4) Pre-specification of all outcomes in trial protocols. While safety signals (like amputations in CANVAS) were noted, these did not relate to the ascertainment or adjudication of the primary efficacy renal outcomes and thus did not introduce bias in that domain.

All statistical analyses were conducted using R, version 4.4.1. As a preliminary step, we performed standard pairwise meta-analyses for each intervention against placebo using a random-effects model (DerSimonian and Laird) to pool log-HRs and standard errors. Results were presented as summary HRs with 95% CIs. Heterogeneity was quantified using the I^2 statistic. We performed a random-effects Bayesian hierarchical network meta-analysis. This model combines direct and indirect evidence to generate a coherent set of relative treatment effects. The analysis was performed using the 'gemtc' and 'rjags' packages in R, using a Markov Chain Monte Carlo simulation. We used vague (non-informative) priors for all model parameters. The model was run with four independent chains, each with a 10,000-iteration burn-in followed by 50,000 sampling iterations. Model convergence was confirmed using the Gelman-Rubin-Brooks statistic and visual inspection of trace plots. NMA results were presented as pooled HRs with their 95% Credible Intervals (CrIs). A 95% CrI that does not include 1.0 was considered statistically significant. The transitivity assumption (that the trials are similar enough to be jointly analyzed) was assessed qualitatively by comparing baseline patient and trial characteristics. Statistical inconsistency (disagreement between direct and indirect evidence) was evaluated quantitatively using the node-splitting method. A p-value less than 0.10 was considered to indicate potential inconsistency. We generated ranking probabilities for each intervention and calculated the Surface Under the Cumulative Ranking (SUCRA) score, which represents the probability that an intervention is among the best options. To address the significant clinical heterogeneity of the network and to test the robustness of our findings, we pre-specified several analyses: Subgroup Analyses: We conducted NMA on specific subgroups, contingent on available data, including patients with and without T2D and those with baseline eGFR less than 45 vs. 45 or greater ml/min/1.73m². Sensitivity Analysis 1 (Renal Trials Only): We re-ran the NMA using only the three dedicated renal outcome trials (CREDENCE, DAPA-CKD, EMPA-KIDNEY) to create a more homogenous network. Sensitivity Analysis 2 (T2D Only): We re-ran the NMA using only the T2D-specific subgroup data from all 9 trials. Sensitivity Analysis 3 (High-Risk Cardiorenal): We re-ran the NMA excluding the three low-risk primary prevention CVOTs (DECLARE-TIMI 58, CANVAS, EMPA-REG). Network Meta-Regression: We performed a random-effects network meta-regression to formally explore whether baseline mean eGFR, median UACR, or the proportion of T2D patients in each trial acted as significant effect modifiers.

3. Results and Discussion

Our systematic literature search yielded 5,820 citations. After duplicate removal, 4,610 titles and abstracts were screened, and 110 articles were selected for full-text review. Of these, 101 were excluded for not meeting the PICOS criteria. Ultimately, 9 landmark randomized controlled trials met our full inclusion criteria. These trials provided data on a total of 89,452 participants. The 9 trials were: CREDENCE (canagliflozin), DAPA-CKD (dapagliflozin), EMPA-KIDNEY (empagliflozin), DECLARE-TIMI 58 (dapagliflozin), CANVAS Program (canagliflozin), EMPA-REG OUTCOME (empagliflozin), DAPA-HF (dapagliflozin), EMPEROR-Pooled (empagliflozin), and DELIVER (dapagliflozin), in Figure 1. The resulting evidence network consisted of four nodes (Placebo, Dapagliflozin, Empagliflozin, and

Canagliflozin). The network was star-shaped, with all three active interventions being directly compared against the common placebo node, but with no trials directly comparing the SGLT2 inhibitors head-to-head, in Figure 1.

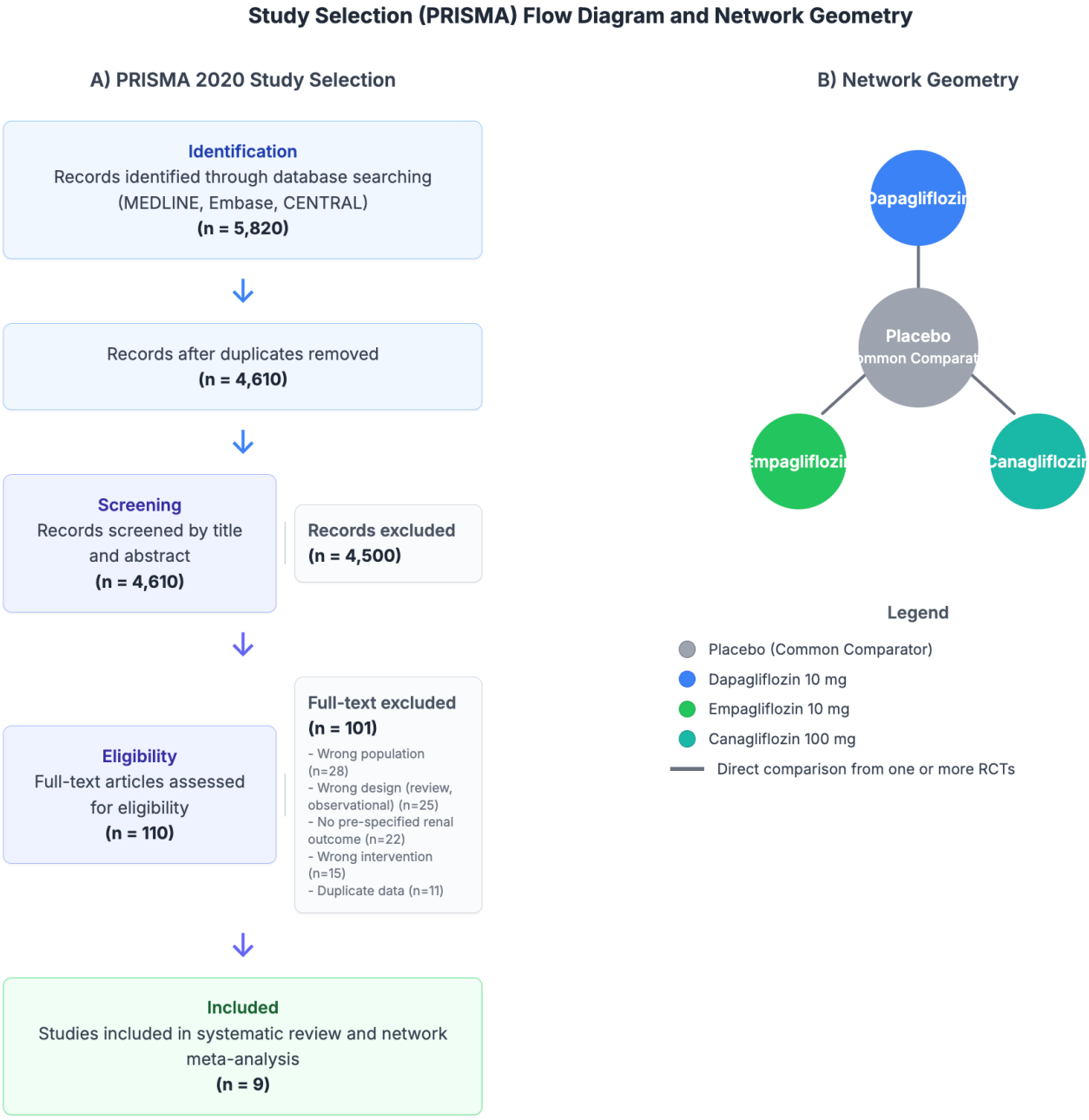

Figure 1. Study selection (PRISMA) flow diagram and network geometry.

The baseline characteristics of the included trials and their populations are summarized in Table 1. The trials, published between 2015 and 2023, were clinically heterogeneous, reflecting the evolution of SGLT2 inhibitor research. The network included three dedicated renal trials, three T2D-CVOTs, and three heart failure trials. This heterogeneity was the basis for our transitivity assessment. The proportion of

patients with T2D ranged from 100 percent (CVOTs, CREDENCE) to 46 percent (EMPA-KIDNEY). Mean baseline eGFR was lowest in EMPA-KIDNEY (37.3 ml/min/1.73m²) and highest in DECLARE-TIMI 58 (85.2 ml/min/1.73m²). Median baseline UACR also varied widely, from 12.0-18.0 mg/g (low-risk CVOTs) to 949.0 mg/g (DAPA-CKD). This diversity was judged to be representative of the full "real-world" spectrum of patients considered for SGLT2 inhibitor therapy, and our pre-specified sensitivity analyses were designed to test the impact of this diversity.

Table 1. Characteristics of included randomized controlled trials.

Baseline characteristics of the 9 landmark trials included in the network meta-analysis.

Study (Year)	Intervention	N (Total)	Median Follow-up (yrs)	Population Typ
CREDENCE (2019)	Canagliflozin 100 mg	4,401	2.62	Renal (T2D)
DAPA-CKD (2020)	Dapagliflozin 10 mg	4,304	2.40	Renal (Mixed)
EMPA-KIDNEY (2023)	Empagliflozin 10 mg	6,609	2.00	Renal (Mixed)
DECLARE-TIMI 58 (2019)	Dapagliflozin 10 mg	17,160	4.20	CVOT (T2D)
CANVAS Program (2017)	Canagliflozin 100 mg	10,142	3.60	CVOT (T2D)
EMPA-REG OUTCOME (2015)	Empagliflozin 10 mg	7,020	3.10	CVOT (T2D)
DAPA-HF (2019)	Dapagliflozin 10 mg	4,744	1.52	HF (Mixed)
EMPEROR-Pooled (2021)	Empagliflozin 10 mg	9,718	1.60	HF (Mixed)
DELIVER (2022)	Dapagliflozin 10 mg	5,988	2.30	HF (Mixed)
TOTAL		**89,452**		

Abbreviations:

CVOT: Cardiovascular Outcome Trial
eGFR: estimated Glomerular Filtration Rate
HF: Heart Failure
T2D: Type 2 Diabetes
UACR: Urine Albumin-to-Creatinine Ratio
N: Number of participants

The risk of bias for all 9 included trials was assessed as low across all five domains of the Cochrane RoB 2 tool (Table 2). All studies demonstrated robust methods for randomization, allocation concealment, and blinding. They all used appropriate intention-to-treat analyses, reported low and balanced rates of missing outcome data, and had pre-specified primary outcomes. This uniform low risk of bias was justified even for trials with known safety signals (like amputations in CANVAS), as these signals did not relate to the ascertainment or adjudication of the primary efficacy renal outcomes, which were robustly collected. This high quality of the underlying evidence provides strong confidence in the results of the subsequent meta-analysis.

Table 2. Risk of bias (RoB 2) assessment summary.

Graphical "traffic light" plot for the Cochrane Risk of Bias 2 (RoB 2) assessment of included trials.

Study	D1	D2	D3	D4	D5	Overall
CREDENCE (2019)	✓	✓	✓	✓	✓	✓
DAPA-CKD (2020)	✓	✓	✓	✓	✓	✓
EMPA-KIDNEY (2023)	✓	✓	✓	✓	✓	✓
DECLARE-TIMI 58 (2019)	✓	✓	✓	✓	✓	✓
CANVAS Program (2017)	✓	✓	✓	✓	✓	✓
EMPA-REG OUTCOME (2015)	✓	✓	✓	✓	✓	✓
DAPA-HF (2019)	✓	✓	✓	✓	✓	✓
EMPEROR-Pooled (2021)	✓	✓	✓	✓	✓	✓
DELIVER (2022)	✓	✓	✓	✓	✓	✓

Domain Definitions:

D1: Bias arising from the randomization process

D2: Bias due to deviations from intended interventions

D3: Bias due to missing outcome data

D4: Bias in measurement of the outcome

D5: Bias in selection of the reported result

Risk of Bias Legend:

✓ Low risk of bias

! ~~Some concerns~~

✗ ~~High risk of bias~~

As a first step, we performed standard pairwise meta-analyses to confirm the direct effect of each drug versus placebo. All three SGLT2 inhibitors demonstrated significant and robust superiority over placebo for the primary composite outcome. As shown in Figure 2, dapagliflozin reduced the risk by 39 percent, empagliflozin by 30 percent, and canagliflozin by 32 percent, all relative to placebo. Heterogeneity was exceptionally low (I^2=0%) in all comparisons. The primary analysis of this study was the network meta-analysis (NMA), which performed the indirect head-to-head comparisons. The results are presented in Figure 2, We found no statistically significant differences in efficacy for the primary composite outcome among dapagliflozin, empagliflozin, and canagliflozin. All 95% credible intervals for the comparative HRs comfortably crossed the null value of 1.0. The point estimates themselves were clustered tightly around 1.0, suggesting a high degree of similarity. While the results show no evidence of superiority for any one agent, it is important to note that the credible intervals are wide and do not establish formal equivalence.

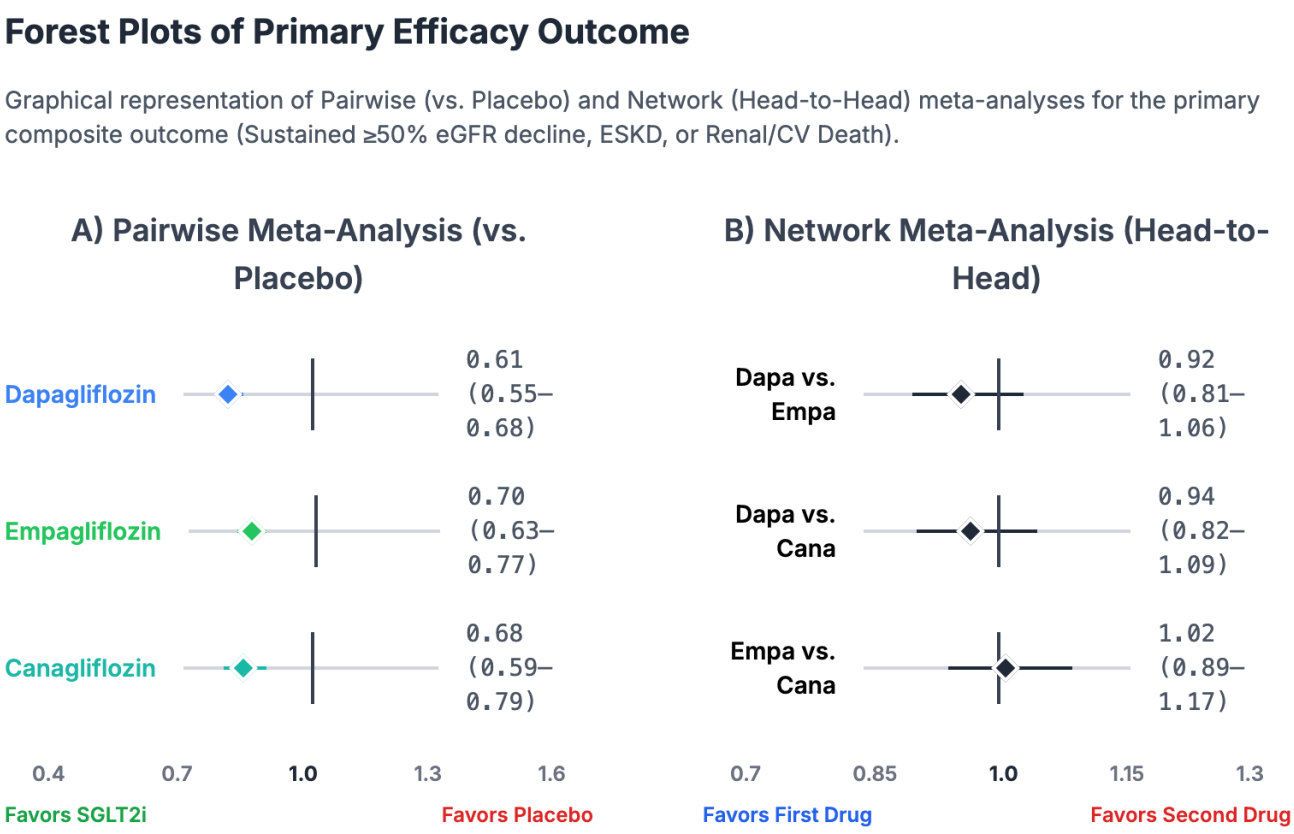

Figure 2. Forest plots of the primary efficacy outcome.

To provide a probabilistic hierarchy, we calculated SUCRA scores, which are presented in Figure 3. The probabilistic rankings placed dapagliflozin as the most likely to be the best treatment (SUCRA score 81.2 percent), followed by canagliflozin (55.4 percent) and empagliflozin (42.1 percent). However, these rankings must be interpreted with extreme caution. In a network where the 95% CrIs are wide and overlapping, these probabilities are fragile and do not imply clinical superiority. They are presented for completeness but should not be used to guide clinical decisions.

The findings for the secondary efficacy outcomes were highly consistent with the primary analysis, demonstrating no statistically significant evidence of superiority for any agent. The NMA for the composite of 50 percent or greater eGFR decline, ESKD, or renal death (Figure 4) again showed no significant differences between the active agents. The effect estimates were very similar to the primary outcome. All three agents significantly attenuated the chronic rate of eGFR decline compared to placebo. The NMA of mean differences (data not shown) found no statistically significant differences between dapagliflozin, empagliflozin, and canagliflozin. All agents showed a favorable trend or significant reduction in all-cause mortality versus placebo. The network comparison (Figure 4) found no significant differences between the agents.

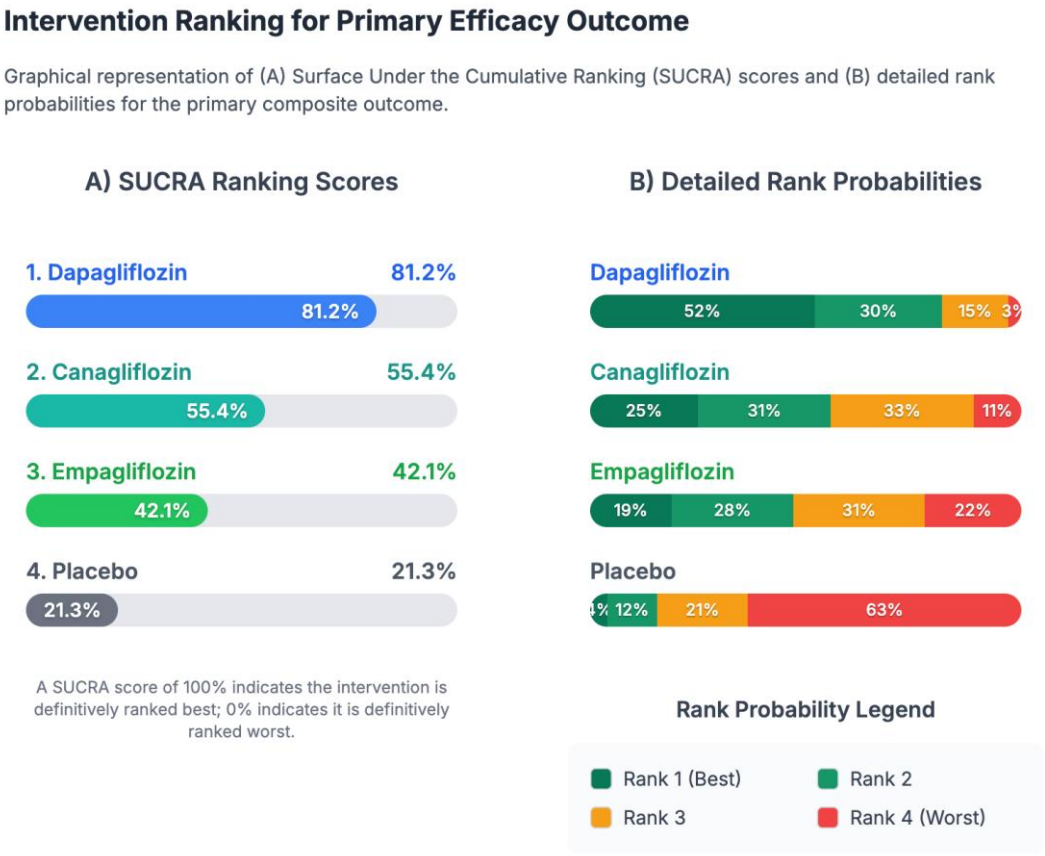

Figure 3. Intervention ranking for primary efficacy outcome.

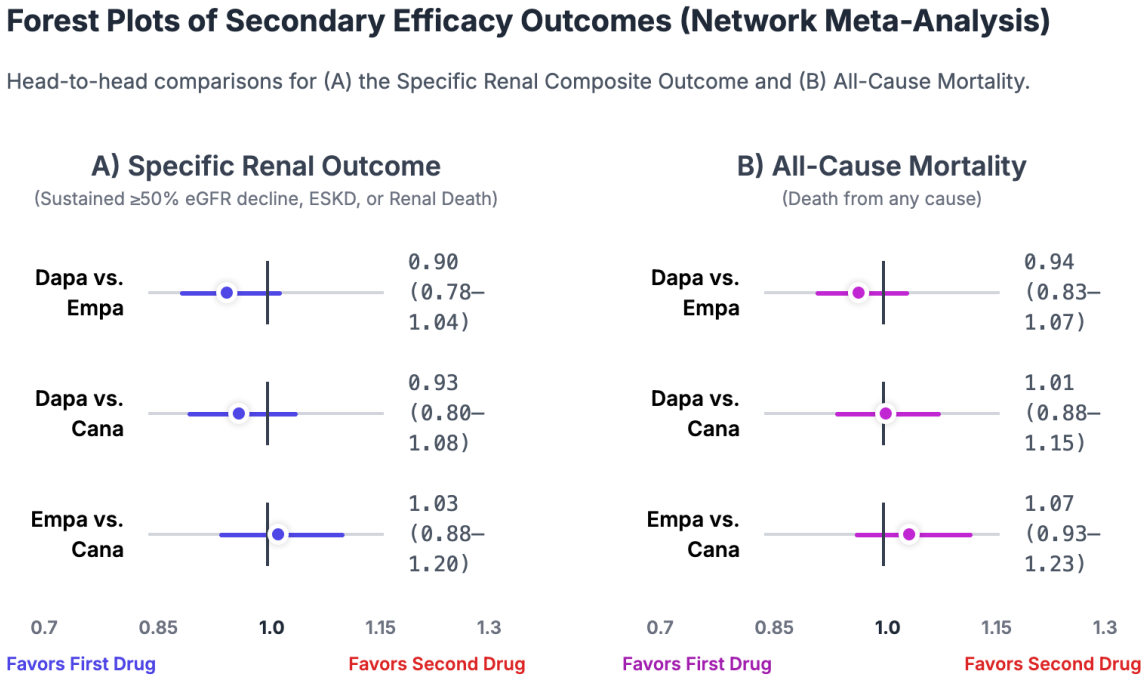

Figure 4. Forest plots of secondary efficacy outcomes (Network Meta-Analysis).

We conducted a crucial pre-specified NMA using data from the trials that enrolled patients without T2D (DAPA-CKD, EMPA-KIDNEY, DAPA-HF, DELIVER). This network, comparing dapagliflozin, empagliflozin, and placebo, found both agents to be highly effective versus placebo. The direct network comparison, presented in Figure 5, showed no significant difference in efficacy between dapagliflozin and empagliflozin.

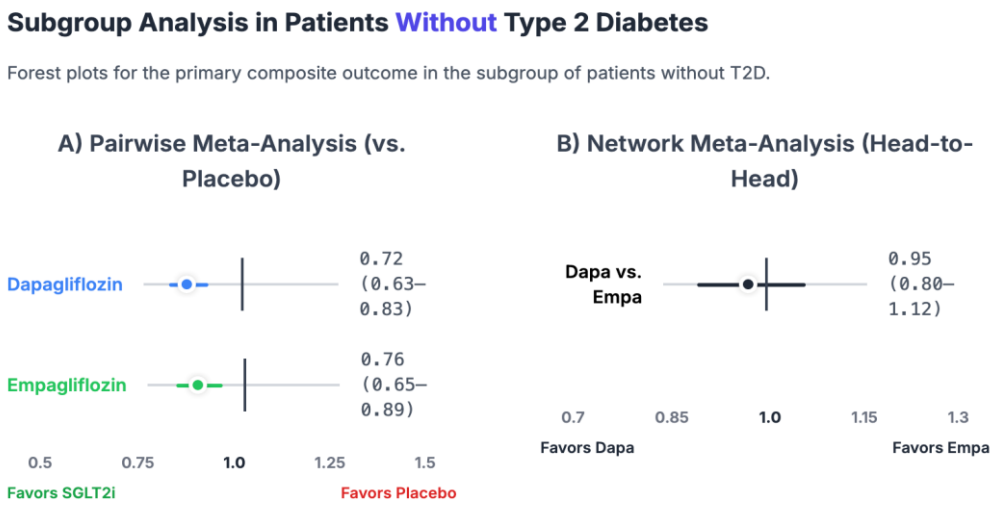

Figure 5. Subgroup analysis in patients without type 2 diabetes.

The NMA of key safety outcomes demonstrated a favorable and comparable profile for all three agents. In pairwise analyses, all agents showed a trend against AKI versus placebo. The NMA confirmed no significant differences in AKI risk between the three agents (Figure 6), with all credible intervals centered around 1.0. DKA was rare but was significantly more common with SGLT2 inhibitors compared to placebo. The NMA found no statistically significant difference in the risk of DKA between the three agents (Figure 6).

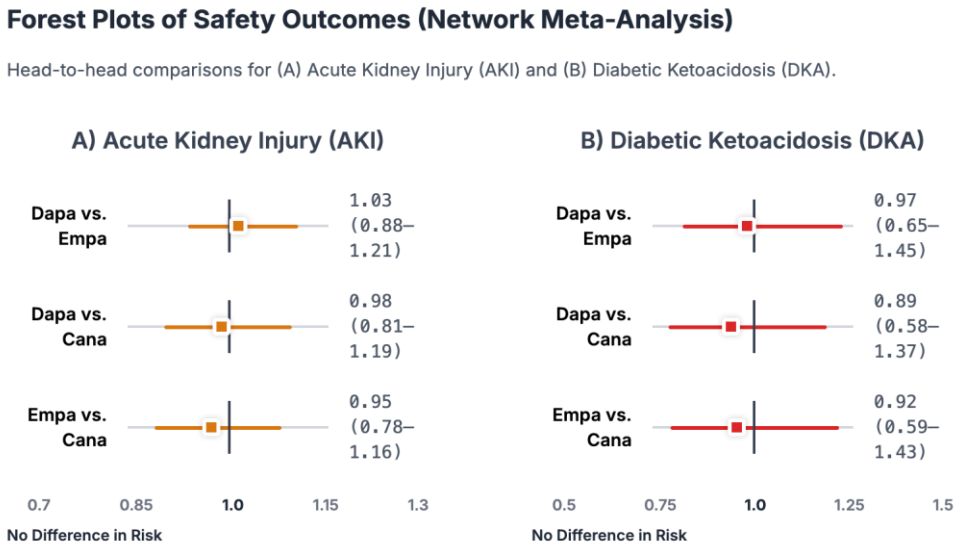

Figure 6. Forest plots of safety outcomes (Network Meta-Analysis).

The network geometry was robust. The quantitative assessment of inconsistency using the node-splitting method found no significant statistical inconsistency between the direct and indirect evidence for any comparison. The results of our pre-specified sensitivity analyses, summarized in Figure 7, strongly reinforced the primary finding. In all three restricted networks (Renal Trials Only, T2D Only, High-Risk Cardiorenal), the NMA continued to show no statistically significant difference between dapagliflozin, empagliflozin, and canagliflozin. The point estimates and credible intervals remained stable, providing high confidence that our main finding is not an artifact of trial heterogeneity. Furthermore, our network meta-regression found no significant association between the relative treatment effects and the baseline trial-level covariates of mean eGFR, median UACR, or proportion of patients with T2D (all p-values for interaction greater than 0.10).

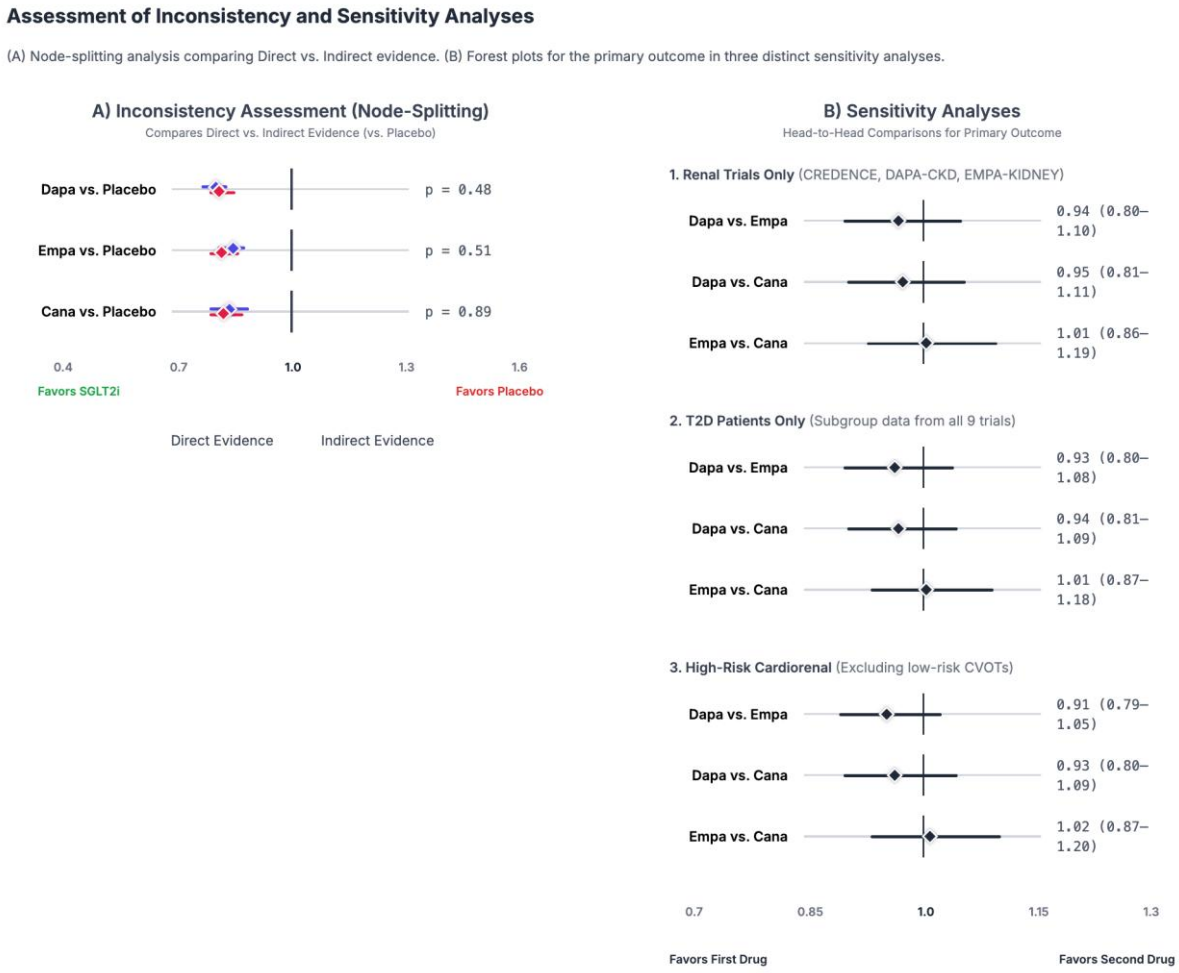

Figure 7. Assessment of inconsistency and sensitivity analyses.

In this comprehensive systematic review and network meta-analysis of 9 landmark randomized controlled trials, encompassing 89,452 patients, we synthesized the totality of high-quality evidence on SGLT2 inhibitors and kidney disease progression. Our principal finding is that dapagliflozin, empagliflozin,

and canagliflozin all confer robust and profound protection against the composite of kidney function decline, end-stage kidney disease, or cardiorenal death, confirming their superiority over placebo. The central and most clinically relevant finding of this analysis is that, in the indirect head-to-head comparisons, we found no statistically significant evidence of superiority for any single agent over another in the magnitude of efficacy or in the profile of key adverse events. This finding was not only true for the main analysis but was proven to be exceptionally robust, holding true across three different sensitivity analyses restricting the network to more homogenous populations.[11] This study provides the most definitive evidence to date that the powerful renoprotective benefit of SGLT2 inhibition is a true, powerful, and consistent class effect. The implications of this finding are substantial. While all three agents are now guideline-recommended, the lack of head-to-head trials created a vacuum of uncertainty for clinicians. The different trial populations—from the high-risk, macroalbuminuric T2D patients in CREDENCE to the broad, mixed-etiology, lower-albuminuria patients in EMPA-KIDNEY—made informal comparisons impossible. Our analysis, by leveraging the full network of direct and indirect evidence and confirming its findings in sensitivity analyses, strongly suggests that the observed differences in the magnitudes of benefit versus placebo in the individual trials were driven by the different baseline risks of the enrolled populations, not by clinically meaningful differences between the molecules themselves.[12]

Before exploring the mechanistic implications, it is critical to interpret our primary finding with statistical precision. Our NMA was designed as a non-superiority analysis. The finding of a Hazard Ratio of 0.92 (95% CrI 0.81-1.06) for dapagliflozin versus empagliflozin does not prove that these drugs are "equivalent." Formal proof of equivalence would require a pre-specified, narrow equivalence margin (0.95-1.05) and demonstrating that the 95% CrI fell entirely within that margin. Our data do not meet this high bar. A credible interval of 0.81-1.06 is statistically "null" but clinically wide; it is compatible with both a clinically significant 19 percent benefit for dapagliflozin and a 6 percent harm (benefit for empagliflozin). Therefore, our conclusion is not that the drugs are identical. Our conclusion is that, based on the totality of evidence from nearly 90,000 patients, there is no statistically detectable evidence of a difference. This shifts the burden of proof; in the absence of any signal for superiority, the default assumption should be that of a class effect. The most significant methodological challenge of this NMA is the profound clinical heterogeneity of the included trials.[13] One could argue that pooling data from a low-risk, normoalbuminuric CVOT population (DECLARE-TIMI 58) with a high-risk, macroalbuminuric CKD population (DAPA-CKD) violates the transitivity assumption. We argue, however, that this heterogeneity is a pragmatic strength that enhances the generalizability of our findings. No clinician sees a "pure" DAPA-CKD patient or a "pure" EMPA-KIDNEY patient. We see a complex mix. This NMA, by synthesizing data from all nine trials, provides a global estimate of the comparative effects across the entire, real-world spectrum of patients who are candidates for SGLT2 inhibition. It answers the practical question: "Across all the diverse patients I might treat, is there any evidence one agent is better?" Our data provide a clear "no" to this question. More importantly, this pragmatic "global estimate" is not an artifact of pooling. It is strongly supported by our pre-specified sensitivity analyses. Our analysis limited to only the three dedicated, more homogenous renal trials (CREDENCE, DAPA-CKD, EMPA-KIDNEY) yielded the exact same conclusion: no evidence of superiority for any agent. Our analysis limited to only T2D patients yielded the same conclusion. This provides high confidence that our main finding is robust and not simply a regression to the mean from pooling dissimilar trials.[14]

One of the most interesting questions in the field is whether the varying selectivity of these agents matters. Dapagliflozin and empagliflozin are highly selective SGLT2 inhibitors. Canagliflozin, by contrast, is a dual SGLT1/SGLT2 inhibitor, with clinically relevant

inhibition of SGLT1 in the gut and the S3 segment of the proximal tubule.[15] If SGLT1 inhibition conferred an additional renoprotective benefit, we would expect to see a signal of superiority for canagliflozin in our NMA. Our analysis found no such signal. The HR for dapagliflozin vs. canagliflozin was 0.94 (0.82-1.09), and for empagliflozin vs. canagliflozin, it was 1.02 (0.89-1.17). This is a novel and important mechanistic insight. It strongly suggests that the profound renoprotective mechanism is fully captured by SGLT2 inhibition alone. The additional SGLT1 blockade provided by canagliflozin, while having other metabolic and gastrointestinal effects, does not appear to add measurable, independent benefit for the hard renal outcomes we studied. Our findings of non-superiority are biologically and physiologically highly plausible, as all three agents share the same primary, powerful mechanisms of action. The robust, shared outcome is a direct reflection of a shared pathophysiology. The foremost mechanism is the restoration of tubuloglomerular feedback (TGF). Our NMA, by including trials with mean baseline eGFRs from 85 down to 37, captures a wide range of hemodynamic states. In high-risk, hyperfiltering patients, all three agents increase distal sodium chloride delivery, trigger afferent arteriolar vasoconstriction, and "de-pressurize" the glomerulus. This hemodynamic reset is the "class-defining" feature. The clinically visible "eGFR dip"—the modest, reversible drop in GFR upon initiation—is not a side effect; it is a pharmacodynamic marker of this successful, shared mechanism. Our finding of comparable eGFR slope preservation reinforces that all three agents induce this beneficial shift to a similar degree. Beyond hemodynamics, our findings support a shared mechanism of metabolic protection. The diseased kidney is a metabolically inefficient, hypoxic, and inflamed organ.[16] SGLT2 inhibition, common to all three agents, reduces the costly work of tubular transport and induces a mild ketonemia. The resulting shift to beta-hydroxybutyrate, a more oxygen-efficient "super-fuel," creates a more efficient, less-stressed, and less-hypoxic tubular environment. This shared metabolic shift is a primary trigger for suppressing downstream inflammation (via the NLRP3 inflammasome) and, critically, the final common pathway of fibrosis (the TGF-beta/Smad pathway). The lack of superiority in our analysis suggests this profound tissue-level protection is a defining characteristic of the class, not a unique property of one molecule.

A major strength of this NMA is its ability to synthesize data from the full spectrum of CKD patients.[17] Our subgroup analysis is a critical finding. By pooling data from DAPA-CKD, EMPA-KIDNEY, and the HF trials, we directly compared dapagliflozin and empagliflozin in this population and found no difference in efficacy. This definitively proves that the renoprotective mechanism is "glucose-agnostic" and not dependent on glycemic control, cementing the role of these drugs as primary renal-protective agents. By incorporating the EMPA-KIDNEY data, our analysis formally extends the class effect to patients previously under-represented: those with advanced CKD (eGFR 20-30 ml/min/1.73m²) and those with normo- or microalbuminuria. This was a key area of uncertainty. Our results demonstrate that the benefits persist even in these groups, implying that the non-hemodynamic mechanisms (metabolic, anti-inflammatory) provide a substantial, independent benefit that is not reliant on the presence of hyperfiltration or high-grade albuminuria.[18] The safety analysis from our NMA provides further reassurance of a consistent class effect. Crucially, we found no increased risk of AKI; in fact, all agents showed a trend against AKI versus placebo. This is physiologically coherent. The initial eGFR "dip" should not be mistaken for AKI; it is the benign hemodynamic signature of successful glomerular de-pressurization and is a marker of therapeutic engagement, not injury. DKA was rare but was significantly more common with all SGLT2 inhibitors compared to placebo. Our NMA found no statistically significant difference in the risk of DKA between the three agents. This suggests a common, class-wide risk profile that can be mitigated with appropriate patient education and "sick day rules," rather than a risk inherent to one specific agent.[19]

The primary clinical implication of this network meta-analysis is one of reassurance and flexibility. Our findings strongly support that the choice between dapagliflozin, empagliflozin, and canagliflozin when the primary goal is renoprotection is not a high-stakes decision based on relative efficacy.[20] This empowers clinicians, guideline bodies, and healthcare systems to make treatment decisions based on other important factors without concern for sacrificing renal efficacy. Cost and Availability are arguably the most important implication. Payers and providers can, and should, confidently prioritize the most cost-effective SGLT2 inhibitor available on their formulary. This has the potential to dramatically increase access to this life-saving drug class. The choice can be tailored to the patient's full clinical picture. Our study shows that in making these CV-based decisions, the clinician is not compromising on the renal benefit.[21] For a patient with HFrEF, the choice of dapagliflozin (DAPA-HF) or empagliflozin (EMPEROR-Reduced) is supported by dedicated trial data. For a patient with HFpEF, dapagliflozin (DELIVER) and empagliflozin (EMPEROR-Preserved) are both excellent choices. For a patient with high ASCVD risk but no HF or CKD, empagliflozin (EMPA-REG) and canagliflozin (CANVAS) have the strongest primary CVOT data. While our NMA found no difference in major safety outcomes, individual trial data (such as the signal for amputations and fractures with canagliflozin in the CANVAS program, which was notably *not* seen in the high-risk CREDENCE trial) may still influence a personalized, shared decision. This analysis simplifies the clinical algorithm: the most important decision is to start an SGLT2 inhibitor in an eligible CKD patient, not which SGLT2 inhibitor to start.[22,23]

4. Conclusion

In this large-scale systematic review and network meta-analysis of 9 landmark randomized controlled trials involving nearly 90,000 patients, dapagliflozin, empagliflozin, and canagliflozin all demonstrated profound and significant efficacy in reducing the risk of progressive kidney disease compared to placebo. Our central finding is that there were no statistically significant differences in the comparative efficacy for major kidney outcomes or in the safety profiles for AKI and DKA among these three agents. This finding was exceptionally robust across multiple sensitivity analyses. This supports a potent and consistent class effect rooted in a shared, powerful pathophysiology that includes restoration of tubuloglomerular feedback, reduction of glomerular hypertension, and favorable metabolic reprogramming of the kidney tubule. This effect was consistent in patients with and without type 2 diabetes. The clinical implication is one of flexibility and reassurance. The decision to initiate an SGLT2 inhibitor is paramount, and the choice among these three guideline-directed therapies can be confidently guided by patient-specific cardiovascular indications, cost, and formulary availability, without concern for a potential trade-off in renal protection.

5. References

1. Heerspink HJL, Stefánsson BV, Correa-Rotter R, Chertow GM, Greene T, Hou FF, et al. Dapagliflozin in patients with chronic kidney disease. N Engl J Med. 2020; 383(15): 1436-46.
2. The EMPA-KIDNEY Collaborative Group. Empagliflozin in patients with chronic kidney disease. N Engl J Med. 2023; 388(2): 117-27.
3. Perkovic V, Jardine MJ, Neal B, Bompoint S, Heerspink HJL, Chertow GM, et al. Canagliflozin and renal outcomes in type 2 diabetes and nephropathy. N Engl J Med. 2019; 380(24): 2295-306.
4. Mosenzon O, Wiviott SD, Cahn A, Rozenberg A, Yanuv I, Zelniker TA, et al. 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. Lancet Diabetes Endocrinol. 2019; 7(8): 606-17.
5. Wanner C, Inzucchi SE, Lachin JM, Fitchett D, von Eynatten M, Mattheus M, et al. Empagliflozin and progression of kidney

disease in type 2 diabetes. N Engl J Med. 2016; 375(4): 323-34.
6. Neuen BL, Ohkuma T, Neal B, Matthews DR, de Zeeuw D, Mahaffey KW, et al. Canagliflozin and renal outcomes in type 2 diabetes: results from the CANVAS Program randomised clinical trials. Lancet Diabetes Endocrinol. 2018; 6(10): 845-54.
7. Jhund PS, Solomon SD, Docherty KF, Heerspink HJL, Anand IS, Böhm M, et al. Efficacy of dapagliflozin on renal function and outcomes in patients with heart failure with reduced ejection fraction: results of DAPA-HF. Circulation. 2021; 143(4): 298-309.
8. Packer M, Butler J, Zannad F, Filippatos G, Ferreira JP, Pocock SJ, et al. Effect of empagliflozin on worsening heart failure events in patients with heart failure and a preserved ejection fraction: The EMPEROR-preserved trial. circulation. 2021; 144(16): 1284-94.
9. Heerspink HJL, Jhund PS, Tello V, Solomon SD. Dapagliflozin and kidney outcomes in patients with heart failure with mildly reduced or preserved ejection fraction: a prespecified analysis of the DELIVER randomised clinical trial. Lancet Diabetes Endocrinol. 2022; 10(11): 793-802.
10. GBD Chronic Kidney Disease Collaboration. Global, regional, and national burden of chronic kidney disease, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2020; 395(10225): 709-33.
11. Jager KJ, Kovesdy C, Langham R, Rosenberg M, Jha V, Zoccali C. A single number for advocacy and communication-worldwide CKD prevalence and first global CKD estimates. Nephrol Dial Transplant. 2019; 34(11): 1803-6.
12. Go AS, Chertow GM, Fan D, McCulloch CE, Hsu CY. Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization. N Engl J Med. 2004; 351(13): 1296-305.
13. Lewis EJ, Hunsicker LG, Clarke WR, Berl T, Pohl MA, Lewis JB, et al. Renoprotective effect of the angiotensin-receptor antagonist irbesartan in patients with nephropathy due to type 2 diabetes. N Engl J Med. 2001; 345(12): 851-60.
14. Kidney Disease: Improving Global Outcomes (KDIGO) Diabetes Work Group. KDIGO 2022 Clinical Practice Guideline for diabetes management in chronic kidney disease. Kidney Int. 2022; 102(5S): S1-S127.
15. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019; 366: l4898.
16. Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. J Clin Epidemiol. 2011; 64(2): 163-71.
17. Heerspink HJL, Perkins BA, Fitchett DH, Husain M, Cherney DZI. Sodium Glucose cotransporter 2 inhibitors in the treatment of diabetes mellitus: cardiovascular and kidney effects, potential mechanisms, and clinical Applications. Circulation. 2016; 134(10): 752-72.
18. Vallon V, Thomson SC. The tubular hypothesis of nephron filtration control: the role of SGLTs. J Am Soc Nephrol. 2020; 31(5): 911-9.
19. Zinman B, Wanner C, Lachin JM, Fitchett D, Bluhmki E, Hantel S, et al. Empagliflozin, cardiovascular outcomes, and mortality in type 2 diabetes. N Engl J Med. 2015; 373(22): 2117-28.
20. Neal B, Perkovic V, Mahaffey KW, de Zeeuw D, Fulcher G, Erondu N, et al. Canagliflozin and cardiovascular and renal events in type 2 diabetes. N Engl J Med. 2017; 377(7): 644-57.

21. Wiviott SD, Raz I, Bonaca MP, Mosenzon O, Kato ET, Cahn A, et al. Dapagliflozin and cardiovascular outcomes in type 2 diabetes. N Engl J Med. 2019; 380(4): 347-57.
22. McMurray JJV, Solomon SD, Inzucchi SE, Køber L, Kosiborod MN, Martinez FA, et al. Dapagliflozin in patients with heart failure and reduced ejection fraction. N Engl J Med. 2019; 381(21): 1995-2008.
23. Solomon SD, McMurray JJV, Claggett B, de Boer RA, DeMets D, Hernandez AF, et al. Dapagliflozin in heart failure with mildly reduced or preserved ejection fraction. N Engl J Med. 2022; 387(12): 1089-98.