







# A hierarchical kidney outcome using win statistics in patients with heart failure from the DAPA-HF and DELIVER trials

Received: 13 November 2023

Accepted: 21 March 2024

Published online: 06 May 2024

 Check for updates

Toru Kondo <sup>1,2</sup>, Pardeep S. Jhund <sup>1</sup>, Samvel B. Gasparian <sup>3</sup>, Mingming Yang <sup>1,4</sup>, Brian L. Claggett<sup>5</sup>, Finnian R. McCausland<sup>6</sup>, Paolo Tolomeo<sup>1</sup>, Muthiah Vadaganathan<sup>5</sup>, Hiddo J. L. Heerspink<sup>7</sup>, Scott D. Solomon <sup>5</sup> & John J. V. McMurray <sup>1</sup> ✉

Win statistics offer a new approach to the analysis of outcomes in clinical trials, allowing the combination of time-to-event and longitudinal measurements and taking into account the clinical importance of the components of composite outcomes, as well as their relative timing. We examined this approach in a post hoc analysis of two trials that compared dapagliflozin to placebo in patients with heart failure and reduced ejection fraction (DAPA-HF) and mildly reduced or preserved ejection fraction (DELIVER). The effect of dapagliflozin on a hierarchical composite kidney outcome was assessed, including the following: (1) all-cause mortality; (2) end-stage kidney disease; (3) a decline in estimated glomerular filtration rate (eGFR) of  $\geq 57\%$ ; (4) a decline in eGFR of  $\geq 50\%$ ; (5) a decline in eGFR of  $\geq 40\%$ ; and (6) participant-level eGFR slope. For this outcome, the win ratio was 1.10 (95% confidence interval (CI) = 1.06–1.15) in the combined dataset, 1.08 (95% CI = 1.01–1.16) in the DAPA-HF trial and 1.12 (95% CI = 1.05–1.18) in the DELIVER trial; that is, dapagliflozin was superior to placebo in both trials. The benefits of treatment were consistent in participants with and without baseline kidney disease, and with and without type 2 diabetes. In heart failure trials, win statistics may provide the statistical power to evaluate the effect of treatments on kidney as well as cardiovascular outcomes.

In patients with heart failure, kidney function is a powerful independent predictor of future heart failure hospitalization and death, irrespective of left ventricular ejection fraction (LVEF)<sup>1–4</sup>. The natural history of heart failure is characterized by progressive worsening of the syndrome over time and this usually includes worsening of kidney function<sup>3,5–7</sup>. Kidney function also influences whether life-saving

pharmacological treatments, including renin–angiotensin system blockers and mineralocorticoid receptor antagonists (MRAs), can be initiated and continued in patients with heart failure and it determines eligibility for transplantation and mechanical circulatory support<sup>8–18</sup>. It is therefore important to understand the effect that new therapies for heart failure have on kidney function; an aspiration with any

<sup>1</sup>British Heart Foundation Cardiovascular Research Centre, University of Glasgow, Glasgow, UK. <sup>2</sup>Department of Cardiology, Nagoya University Graduate School of Medicine, Nagoya, Japan. <sup>3</sup>Late-Stage Development, Cardiovascular, Renal, and Metabolism, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden. <sup>4</sup>Department of Cardiology, Zhongda Hospital, School of Medicine, Southeast University, Nanjing, China. <sup>5</sup>Cardiovascular Division, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA. <sup>6</sup>Renal Division, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA. <sup>7</sup>Department of Clinical Pharmacy and Pharmacology, University of Groningen, Groningen, the Netherlands.

✉ e-mail: [john.mcmurray@glasgow.ac.uk](mailto:john.mcmurray@glasgow.ac.uk)

**Table 1 | Participant characteristics in the pooled DAPA-HF and DELIVER dataset**

	Dapagliflozin n=5,503	Placebo n=5,501	P
Age, years	69.4±10.6	69.4±10.4	0.98
Female sex	1,928 (35.0)	1,927 (35.0)	>0.99
Region			0.96
Europe and Saudi Arabia	2,588 (47.0)	2,571 (46.7)	
North America	762 (13.8)	764 (13.9)	
South America	1,003 (18.2)	994 (18.1)	
Asia and the Pacific	1,150 (20.9)	1,172 (21.3)	
Ethnicity			0.40
White	3,875 (70.4)	3,894 (70.8)	
Asian	1,182 (21.5)	1,208 (22.0)	
Black or African American	203 (3.7)	182 (3.3)	
Other	243 (4.4)	217 (3.9)	
Baseline body mass index, kg m <sup>-2</sup>	29.1±6.1	29.1±6.1	0.99
<b>Vital signs</b>			
Heart rate, beats per min	71.5±11.7	71.5±11.7	0.98
Systolic blood pressure, mmHg	125.6±16.1	125.4±16.1	0.54
Diastolic blood pressure, mmHg	73.8±10.4	73.7±10.5	0.61
<b>Laboratory values</b>			
HbA1c, %	6.6±1.4	6.5±1.4	0.69
Creatinine, μmol l <sup>-1</sup>	103.0±30.6	103.6±31.0	0.28
eGFR, ml min <sup>-1</sup> 1.73 m <sup>-2</sup>	63.2±19.4	62.9±19.4	0.35
eGFR <60 ml min <sup>-1</sup> 1.73 m <sup>-2</sup> (%)	2,478 (45.0)	2,518 (45.8)	0.43
NT-proBNP, ng l <sup>-1</sup>	1,178 (708–2,129)	1,178 (698–2,117)	0.63
NT-proBNP, ng l <sup>-1</sup> if baseline ECG in AF/AFL	1,569 (1,034–2,534)	1,525 (1,033–2,481)	0.49
NT-proBNP, ng l <sup>-1</sup> if baseline ECG not in AF/AFL	970 (563–1,820)	961 (566–1,846)	>0.99
AF/AFL on ECG (%)	1,896 (34.5)	1,875 (34.1)	0.68
<b>Heart failure characteristics</b>			
Previous hospitalization due to heart failure	2,393 (43.5)	2,395 (43.5)	0.96
Time from diagnosis of heart failure			0.29
≤1 year	1,520 (27.6)	1,579 (28.7)	
1–5 years	2,173 (39.5)	2,181 (39.7)	
>5 years	1,807 (32.9)	1,739 (31.6)	
NYHA functional class			0.097
I and II	3,919 (71.2)	3,995 (72.6)	
II and IV	1,584 (28.8)	1,505 (27.4)	
Baseline KCCQ-TSS	75.0 (56.3–89.6)	75.0 (58.3–89.6)	0.40
Baseline LVEF (%)	44.2±13.7	44.2±14.1	0.75
<b>Clinical history</b>			
T2D	2,394 (43.5)	2,395 (43.5)	0.97
AF	2,627 (47.7)	2,655 (48.3)	0.58

**Table 1 (continued) | Participant characteristics in the pooled DAPA-HF and DELIVER dataset**

	Dapagliflozin n=5,503	Placebo n=5,501	P
Hypertension	4,516 (82.1)	4,558 (82.9)	0.27
Myocardial infarction	1,830 (33.3)	1,900 (34.5)	0.15
Stroke	524 (9.5)	539 (9.8)	0.62
<b>Medical therapy</b>			
ACEi	2,476 (45.0)	2,480 (45.1)	0.93
ARB	1,807 (32.8)	1,769 (32.2)	0.45
ARNI	415 (7.5)	394 (7.2)	0.45
Beta-blocker	4,869 (88.5)	4,863 (88.4)	0.90
MRA	3,036 (55.2)	3,000 (54.5)	0.50
Loop diuretic	4,309 (78.3)	4,326 (78.6)	0.67
Digitalis	595 (10.8)	587 (10.7)	0.81
CRT-D or ICD	709 (12.9)	700 (12.7)	0.80

Data are presented as the mean ± s.d. or median (interquartile range) for continuous measures, and n (%) for categorical measures. Continuous variables were compared using a two-sided t-test or Wilcoxon rank-sum test; categorical variables were compared using a chi-squared test. Adjustments for multiple comparisons were not made. Body mass index was missing in eight participants; NT-proBNP was missing in one participant; AF/AFL on ECG was missing in two participants; time from diagnosis of heart failure was missing in five participants; and NYHA functional class was missing in one participant. ACEi, angiotensin-converting enzyme inhibitor; AF, atrial fibrillation; AFL, atrial flutter; ARB, angiotensin receptor blocker; ARNI, angiotensin receptor neprilysin inhibitor; CRT-D, cardiac resynchronization therapy-defibrillator; ECG, electrocardiogram; HbA1c, glycated hemoglobin; ICD, implantable cardioverter-defibrillator; KCCQ-TSS, Kansas City Cardiomyopathy Questionnaire-Total Symptom Score; NT-proBNP, N-terminal pro-B-type natriuretic peptide; NYHA, New York Heart Association.

treatment for heart failure is to at least preserve and, ideally, improve kidney function.

Unfortunately, few trials in patients with heart failure have been large enough and long enough to accrue a sufficient number of 'hard' kidney endpoints to allow a statistically robust evaluation of these outcomes using conventional statistical approaches, for example, time-to-first-occurrence of death, end-stage kidney disease (ESKD) or a large decline in estimated glomerular filtration rate (eGFR)<sup>19–23</sup>. The rate of decline over time (slope) in eGFR has been used as an alternative means of evaluating the effect of treatment on kidney function; however<sup>24–26</sup>, while statistically more powerful, this measure does not incorporate death or initiation of renal replacement therapy and the clinical relevance of small changes in eGFR slope have been questioned.

The use of hierarchical composite endpoints analyzed with win statistics may solve some of these problems by integrating death, relatively infrequent major kidney events (for example, ESKD), the occurrence of large changes in eGFR that are somewhat more frequent, and changes in the eGFR slope, with each of these components ordered in a hierarchy reflecting their clinical importance<sup>27–29</sup>. The hierarchical composite outcome created by this approach consists of components, all of which reflect the progression of kidney disease, and this endpoint is both clinically relevant and statistically powerful<sup>30</sup>.

In this post hoc study, we evaluated the effects of dapagliflozin on kidney function in patients with heart failure and reduced ejection fraction, and heart failure and mildly reduced or preserved ejection fraction<sup>31,32</sup>, using a hierarchical composite kidney outcome, analyzed using win statistics.

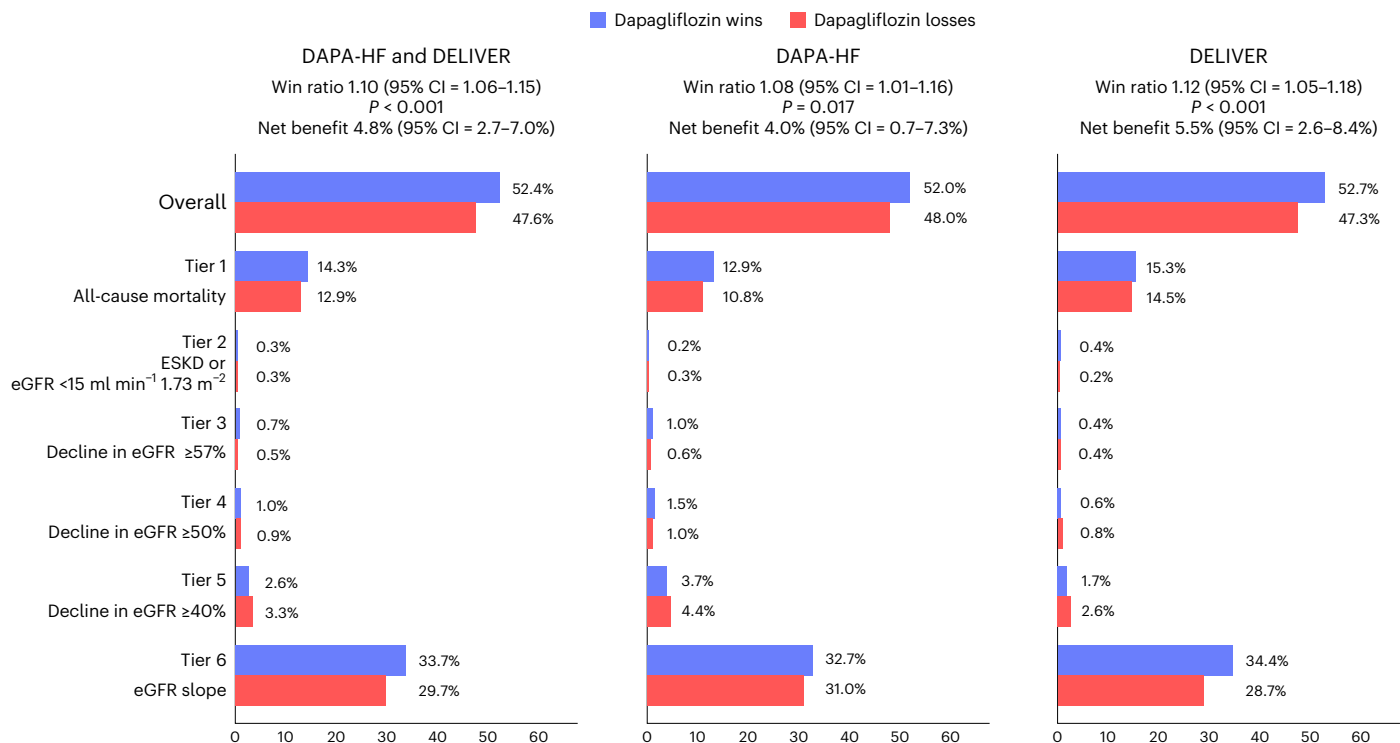
## Results

Of the 11,004 participants included in the Dapagliflozin and Prevention of Adverse Outcomes in Heart Failure (DAPA-HF) and Dapagliflozin Evaluation to Improve the Lives of Patients with Preserved Ejection

**Table 2 | Outcomes analyzed using conventional statistical approaches**

	DAPA-HF and DELIVER						DAPA-HF						DELIVER					
	DAPA (n=5,503)			Placebo (n=5,501)			DAPA (n=2,372)			Placebo (n=2,370)			DAPA (n=3,131)			Placebo (n=3,131)		
	Value	Rate per 100 patient-years (95% CI)	Value	Rate per 100 patient-years (95% CI)	1/HR (95% CI) <sup>a</sup>	Value	Rate per 100 patient-years (95% CI)	Value	Rate per 100 patient-years (95% CI)	1/HR (95% CI) <sup>a</sup>	Value	Rate per 100 patient-years (95% CI)	Value	Rate per 100 patient-years (95% CI)	1/HR (95% CI) <sup>a</sup>			
Composite of all-cause mortality, ESKD or eGFR <15 ml min <sup>-1</sup> .73 m <sup>-2</sup> , decline in eGFR ≥40%, n	1,111	15.7 (14.8–16.7)	1,151	16.2 (15.3–17.2)	1.04 (0.95–1.12)	458	15.0 (13.7–16.4)	509	16.8 (15.4–18.4)	1.12 (0.99–1.27)	653	16.3 (15.1–17.6)	642	15.7 (14.6–17.0)	0.98 (0.88–1.09)			
Composite of all-cause mortality, ESKD or eGFR <15 ml min <sup>-1</sup> .73 m <sup>-2</sup> , decline in eGFR ≥50%, n	884	12.2 (11.4–13.0)	979	13.5 (12.7–14.4)	1.11 (1.02–1.22)	330	10.5 (9.4–11.7)	402	13.0 (11.8–14.3)	1.24 (1.07–1.43)	554	13.5 (12.4–14.7)	577	13.9 (12.8–15.1)	1.04 (0.93–1.17)			
All-cause mortality, n	769	7.5 (7.0–8.0)	847	8.3 (7.7–8.8)	1.11 (1.01–1.22)	275	8.0 (7.1–9.0)	326	9.6 (8.6–10.7)	1.20 (1.02–1.41)	494	7.2 (6.6–7.8)	521	7.6 (7.0–8.3)	1.06 (0.94–1.20)			
ESKD or eGFR <15 ml min <sup>-1</sup> .73 m <sup>-2</sup> , n	34	0.3 (0.2–0.5)	40	0.4 (0.3–0.5)	1.19 (0.75–1.88)	22	0.7 (0.4–1.0)	22	0.7 (0.4–1.0)	1.02 (0.56–1.83)	12	0.2 (0.1–0.3)	18	0.3 (0.2–0.4)	1.51 (0.73–3.13)			
Decline in eGFR ≥57%, n	59	0.8 (0.6–1.0)	70	1.0 (0.8–1.2)	1.20 (0.85–1.69)	27	0.8 (0.6–1.2)	35	1.1 (0.8–1.5)	1.33 (0.80–2.19)	32	0.8 (0.5–1.1)	35	0.8 (0.6–1.2)	1.09 (0.67–1.75)			
Decline in eGFR ≥50%, n (%)	132	1.8 (1.5–2.1)	147	2.0 (1.7–2.4)	1.12 (0.89–1.42)	61	1.9 (1.5–2.5)	83	2.6 (2.1–3.3)	1.39 (1.00–1.94)	71	1.7 (1.4–2.2)	64	1.5 (1.2–2.0)	0.89 (0.64–1.25)			
Decline in eGFR ≥40%, n	388	5.6 (5.1–6.2)	356	5.0 (4.5–5.5)	0.90 (0.78–1.03)	202	6.5 (5.7–7.5)	208	6.8 (5.9–7.8)	1.04 (0.86–1.26)	196	4.9 (4.2–5.6)	148	3.6 (3.1–4.3)	0.75 (0.60–0.92)			
eGFR slope, ml min <sup>-1</sup> .73 m <sup>-2</sup> per year <sup>b</sup>	-1.77±0.07	-	-2.28±0.07	-	-	-2.76±0.11	-	-3.22±0.11	-	-	-1.03±0.08	-	-1.56±0.08	-	-			

<sup>a</sup>Models were stratified according to diabetes status (and according to trial in the pooled dataset). <sup>b</sup>Data are presented as the mean±s.e.



**Fig. 1** Effect of dapagliflozin on the hierarchical composite kidney outcome. Win statistics were two-sided. Models were stratified according to diabetes status (and according to trial in the pooled dataset). Adjustments were not made for multiple comparisons. The exact *P* values were 0.00001 in the pooled dataset and 0.0002 in the DELIVER dataset.

Fraction Heart Failure (DELIVER) trials, 4,742 were enrolled in DAPA-HF and 6,262 in DELIVER. Participants were assigned equally to dapagliflozin ( $n = 5,503$ ) or placebo ( $n = 5,501$ ).

### Participants

The participant characteristics according to the randomized treatment groups were well-balanced at baseline (Table 1). In the pooled dataset, there were 1,111 composite events of all-cause mortality, a decline of  $\geq 40\%$  in eGFR or ESKD or an eGFR  $< 15 \text{ ml min}^{-1} 1.73 \text{ m}^{-2}$ , in the dapagliflozin group, and 1,151 events in the placebo group; in the DAPA-HF trial, there were 458 in the dapagliflozin group and 509 in the placebo group; in the DELIVER trial, there were 653 in the dapagliflozin group and 642 in the placebo group (Table 2). The effects of dapagliflozin on conventional composite outcomes, analyzed as the time-to-first event, are shown in Table 2. In the pooled dataset, the total eGFR slope in the dapagliflozin group was significantly lower than in the placebo group ( $-1.77 \pm 0.07$  (mean  $\pm$  s.e.) versus  $-2.28 \pm 0.07 \text{ ml min}^{-1} 1.73 \text{ m}^{-2}$  per year,  $P < 0.001$ ) (Table 2 and Extended Data Fig. 1). Similarly, in DAPA-HF and DELIVER separately, the total eGFR slope in the dapagliflozin group was significantly less steep than in the placebo group (DAPA-HF,  $-2.76 \pm 0.11$  (mean  $\pm$  s.e.) versus  $-3.22 \pm 0.11 \text{ ml min}^{-1} 1.73 \text{ m}^{-2}$  per year,  $P < 0.001$ ; DELIVER,  $-1.03 \pm 0.08$  (mean  $\pm$  SE) versus  $-1.56 \pm 0.08 \text{ ml min}^{-1} 1.73 \text{ m}^{-2}$  per year,  $P = 0.004$ ).

### Win ratio and proportion of wins and losses in each tier

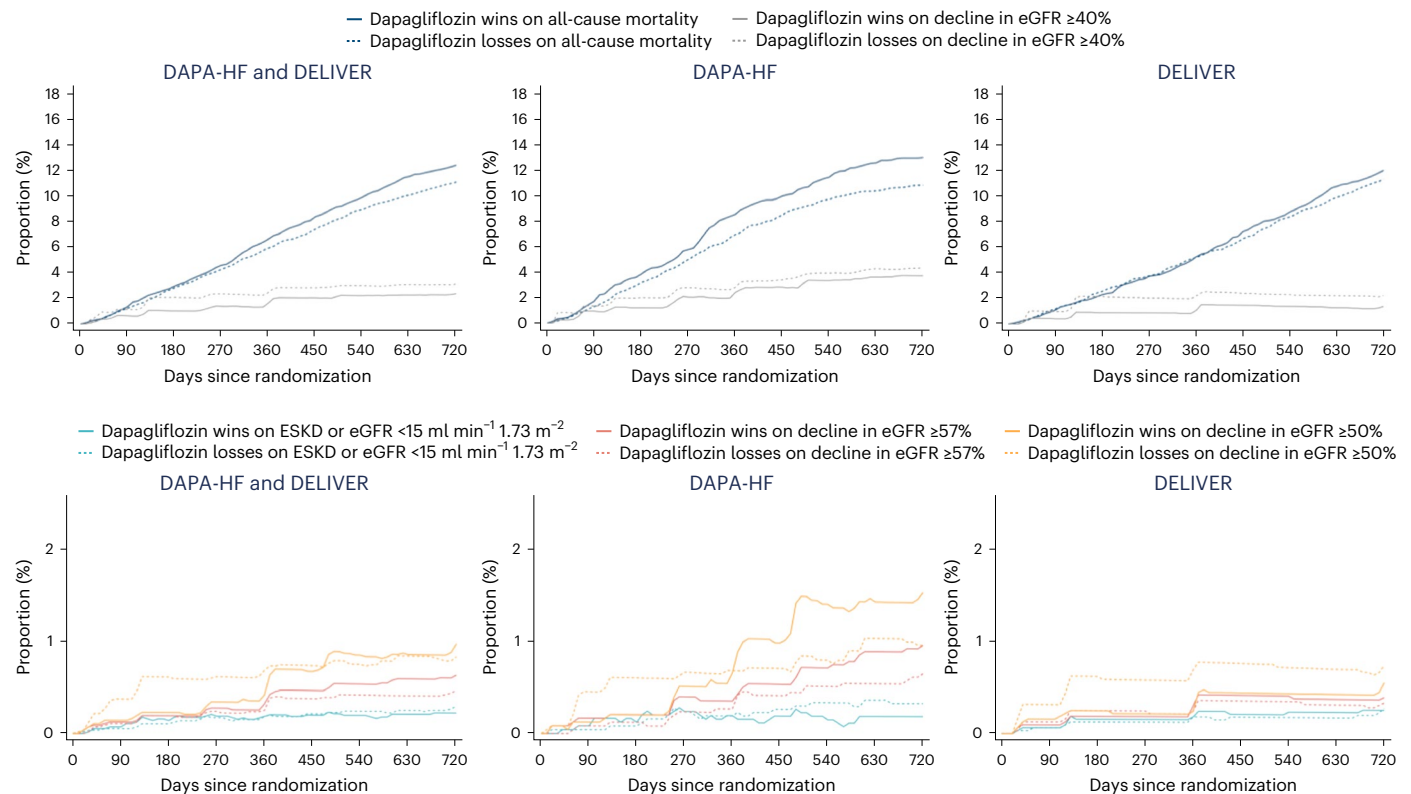
The effects of dapagliflozin on the hierarchical composite kidney outcome, as estimated using win statistics, are summarized in Fig. 1. The hierarchical composite kidney outcome included the following tiers: (1) all-cause mortality; (2) ESKD or eGFR  $< 15 \text{ ml min}^{-1} 1.73 \text{ m}^{-2}$ ; (3) a decline in eGFR of  $\geq 57\%$ ; (4) a decline in eGFR of  $\geq 50\%$ ; (5) a decline in eGFR of  $\geq 40\%$ ; and (6) participant-level eGFR slope. The win ratio was 1.10 (95% confidence interval (CI) = 1.06–1.15) in the pooled dataset, 1.08 (95% CI = 1.01–1.16) in DAPA-HF dataset and 1.12 (95% CI = 1.05–1.18) in the DELIVER dataset, demonstrating that dapagliflozin was superior

to placebo with regard to the hierarchical composite kidney outcome compared in all three analyses. The eGFR slope accounted for most wins and losses, and incorporation of the participant-level eGFR slope in this model reduced the proportion of ties that would have occurred (in 63.4% of pairs in the pooled DAPA-HF and DELIVER dataset). The net benefit was 4.8% (95% CI = 2.7–7.0%) in the pooled dataset, 4.0% (95% CI = 0.7–7.3%) in the DAPA-HF dataset and 5.5% (95% CI = 2.6–8.4%) in the DELIVER dataset.

**Sensitivity analyses.** In the sensitivity Model 1 analysis, which excluded the tier for a decline in eGFR of  $\geq 40\%$ , win ratios remained higher than 1.0 for participants in the pooled dataset, and in the DAPA-HF and DELIVER trials separately (Extended Data Fig. 2). In sensitivity Model 2, which excluded both the tier for a decline in eGFR of  $\geq 40\%$  and the eGFR slope, the lower CIs of the win ratios and win odds (accounting for ties because of the exclusion of the eGFR slope) were not higher than 1.0 in the DELIVER dataset (Extended Data Fig. 3). The win ratios obtained using sensitivity Model 2 were similar to the 1/hazard ratios (HRs) for the composite kidney endpoints estimated using conventional statistical approaches and evaluated with the similar composite of all-cause mortality, ESKD or eGFR  $< 15 \text{ ml min}^{-1} 1.73 \text{ m}^{-2}$ , or decline in eGFR of  $\geq 50\%$  (Table 2). Adding the eGFR slope back into sensitivity Model 2 increased the net benefit from 1.7% to 5.3% in the pooled dataset. In sensitivity Model 3, which excluded all-cause mortality, almost identical results to the main model were observed in the pooled dataset, and the DAPA-HF and DELIVER datasets separately (Extended Data Fig. 4).

### Proportions of wins and losses over time

For all-cause mortality, differences in the proportion of wins and losses between treatments increased gradually over time in the pooled dataset, and in the DAPA-HF and DELIVER datasets separately (Fig. 2). In the three datasets, the proportion of losses with dapagliflozin for a decline in eGFR of  $\geq 40\%$  was larger than that of wins, but this difference narrowed over time. The proportions of wins and losses for ESKD or an



**Fig. 2 | Proportion of wins and losses over time.** Each figure was plotted every 10 days for up to 720 days.

eGFR  $< 15 \text{ ml min}^{-1} 1.73 \text{ m}^{-2}$ , and declines in eGFR of  $\geq 57\%$  and  $\geq 50\%$ , were small and differed little throughout the follow-up. For comparison, the effects of dapagliflozin versus placebo, plotted using the Kaplan–Meier method are shown in Extended Data Fig. 5.

### Win ratio and proportions of wins and losses in the subgroups

Win ratios, and the proportion of wins and losses, in the dapagliflozin groups according to a history of type 2 diabetes (T2D), eGFR category ( $< 60$  versus  $\geq 60 \text{ ml min}^{-1} 1.73 \text{ m}^{-2}$ ) are shown in Fig. 3. The treatment effect estimate from the win ratio analysis was consistent across these subgroups, that is, there were no apparent differences in the estimates.

### Power analysis

When using a hierarchical composite endpoint, sample size requirements are smaller than the time-to-first composite endpoint evaluated using the Cox proportional hazards model (Extended Data Fig. 6).

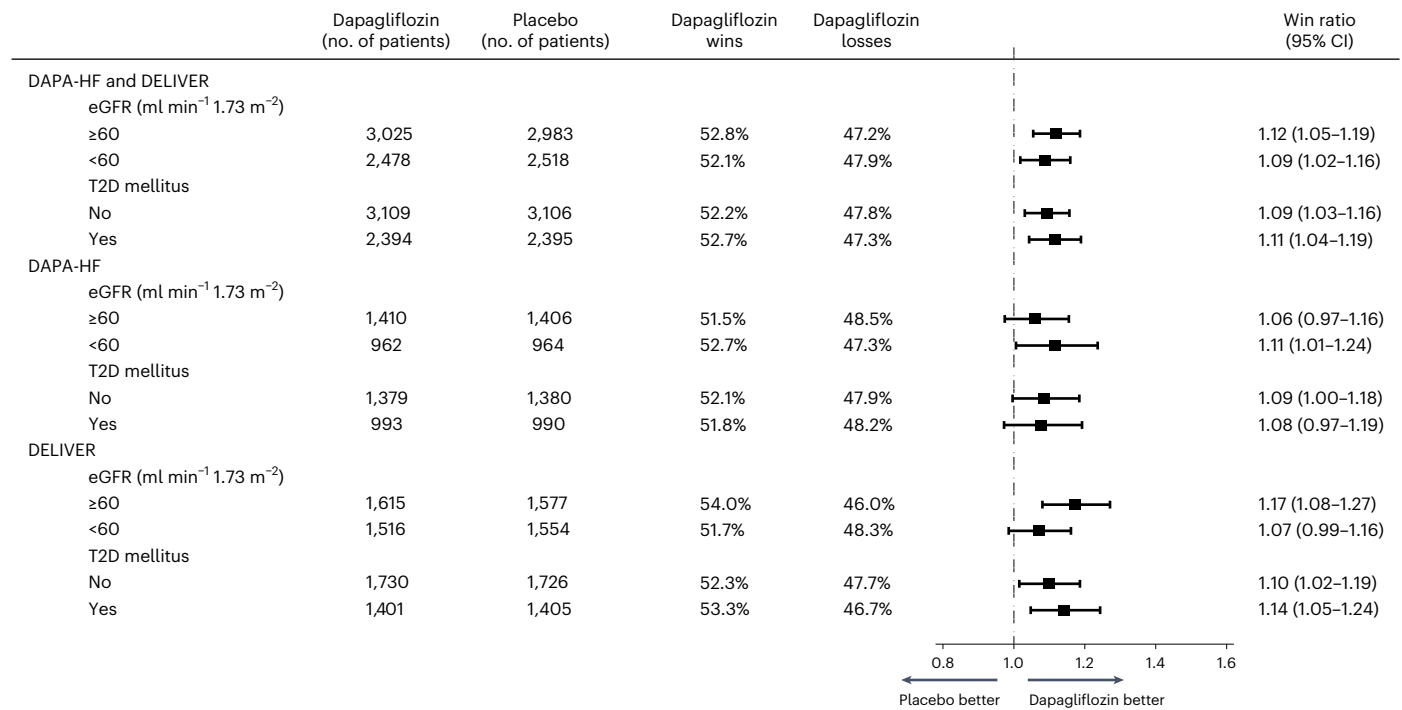
### Discussion

These post hoc analyses show how win statistics can be used to demonstrate the benefit of a treatment for heart failure (in this case, dapagliflozin) on kidney function in patients with both heart failure and reduced ejection fraction and heart failure and mildly reduced or preserved ejection fraction. It is generally difficult to demonstrate the potential kidney benefits of cardiovascular drugs using a conventional renal endpoint because of the small number of events in an ‘unenriched’ population (for example, without albuminuria) during a relatively short-term follow-up. In such a setting, the hierarchical composite endpoint examined in the present study provides greater statistical power and may offer the opportunity to demonstrate both cardiovascular and kidney benefits in the same population in the same trial. In addition to the summary of win statistics usually shown in analyses of this type, we also presented the proportion of wins and losses over time, similar to the depiction of event rates over time provided using traditional statistical methods.

Although superficially similar, the win statistics approach used in this study differs substantially from time-to-first-event analysis for a composite endpoint. The most obvious difference is that events are analyzed according to a hierarchy<sup>27,28</sup>. All-cause mortality was the most significant event in the composite hierarchical outcome and was tested as the first tier in the hierarchy. Unlike time-to-first-event analysis, the win statistics approach includes all deaths, including those occurring after a worsening kidney disease event. With the win statistics approach, a hierarchy of worsening kidney disease events was also created, reflecting their clinical importance, for example, the development of ESKD or an eGFR  $< 15 \text{ ml min}^{-1} 1.73 \text{ m}^{-2}$ , and large decreases in eGFR. As a further refinement, it is also possible to extend the hierarchy to include different proportional declines in eGFR; in the present analysis, we incorporated declines in eGFR of  $\geq 57\%$ ,  $\geq 50\%$  and  $\geq 40\%$ . An additional advantage of win statistics is that the hierarchical composite outcome can logically incorporate continuous variables such as the eGFR slope<sup>27–29</sup>. Because the statistical power for conventional composite kidney outcomes is often insufficient when analyzing events such as those discussed above (because of their low incidence rate in some populations), analysis of the eGFR slope has been suggested as an alternative<sup>19–22,24</sup>. However, the eGFR slope is evaluated as a single ‘stand-alone’ outcome; its interpretation alongside other more important kidney endpoints simultaneously may not be easy. By contrast, the win statistics approach provides an outcome that integrates all relevant outcomes and all patients contribute to the analysis. One issue with the eGFR slope, either as a stand-alone endpoint or part of the win ratio approach, is that some drugs may cause an initial decline in eGFR<sup>33–35</sup>. The slope after initiation may more accurately reflect the chronic effect of these drugs, but may overestimate treatment benefit<sup>36,37</sup>; thus, more appropriately, we calculated the eGFR slope over the whole treatment period using a piece-wise, linear, two-slope model accounting for the effects of the acute and chronic phases<sup>38</sup>.

A closer look at the proportion of wins and losses revealed several findings. Despite the less steep eGFR slope with dapagliflozin compared





**Fig. 3 | Effect of randomized treatments on the hierarchical composite kidney outcome according to selected subgroups.** Models were stratified according to diabetes status (and according to the trial in the pooled dataset). The squares indicate the win ratios and the bars indicate the upper and lower boundaries of the 95% CI.

to placebo, the proportion of wins with dapagliflozin (over placebo) for tier 5 of the hierarchy (that is, a decline in eGFR of  $\geq 40\%$ ) was lower than the proportion of losses. The probable explanation for this is that DAPA-HF and DELIVER did not have an active run-in period and the initial drop in eGFR in some patients randomized to dapagliflozin led to a decline in eGFR counting as an ‘event’<sup>31,32,39–42</sup>. On examining the proportion of wins and losses over time, it can also be seen that the difference in the tier representing a decline in eGFR of  $\geq 40\%$ , which may reflect the initial drop with dapagliflozin early after randomization, was progressively smaller over time in the DAPA-HF and DELIVER, supporting this explanation and identifying the longer-term benefit of dapagliflozin on the kidney. Indeed the kidney benefits of both these drugs were more apparent over time, observed as the changing proportion of wins and losses over time, which is analogous to the divergence of Kaplan–Meier plots using conventional analysis.

Win statistics are a relatively new approach to analyzing trial data and may still be unfamiliar to some physicians<sup>43,44</sup>. However, their use is increasing rapidly, particularly in cardiovascular medicine; several recent trials had primary endpoints analyzed using win statistics<sup>45–52</sup>. At least one treatment has received regulatory approval based on a trial of this type<sup>45</sup>. Next, there is always debate about which components to include in a hierarchical composite outcome and these should be discussed between the relevant stakeholders, including patients, clinical trialists, and regulatory and reimbursement agencies. Although all-cause mortality is usually included as the first tier in such analyses, it could be argued that this is not a kidney-specific outcome<sup>30</sup>. To address this concern, we added a sensitivity analysis excluding all-cause mortality from the hierarchy, which showed essentially the same findings. Third, treatments may not affect each component of a composite outcome equally, although this is also an issue with composite endpoints evaluated using conventional statistics. Therefore, it is important to examine the proportion of wins or losses for each component of the composite to interpret the overall result.

This study has several limitations. eGFR was obtained at different scheduled visits in the two trials, while the incidence of the renal

endpoints defined according to eGFR may have been affected by the frequency of the eGFR measurements. The hierarchical composite renal outcome used in this study was created post hoc. However, the selected hierarchy reflected the natural progression of kidney disease. It was validated in multiple sensitivity models and by comparison with the analysis of a conventional composite outcome analyzed using a standard method. The thresholds for declines in eGFR were also decided post hoc; thus, ‘sustained’ eGFR decline could not be confirmed using repeat measurement. The eGFR slope may also have been affected by the number of scheduled visits, visit intervals and the follow-up period in each trial.

In conclusion, it was possible to create a comprehensive, multicomponent, hierarchical composite kidney endpoint that is both clinically relevant and statistically powerful when analyzed using win statistics. With this approach, we confirmed the benefits of dapagliflozin on kidney function in patients with heart failure. This benefit was observed regardless of LVEF, baseline eGFR and T2D status. This approach can improve the power and precision around the estimate of effects on kidney outcomes and should be considered in future heart failure trials.

### Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41591-024-02941-8>.

### References

- Hillege, H. L. et al. Renal function as a predictor of outcome in a broad spectrum of patients with heart failure. *Circulation* **113**, 671–678 (2006).
- Damman, K. et al. Renal impairment, worsening renal function, and outcome in patients with heart failure: an updated meta-analysis. *Eur. Heart J.* **35**, 455–469 (2014).

3. Damman, K. & Testani, J. M. The kidney in heart failure: an update. *Eur. Heart J.* **36**, 1437–1444 (2015).
4. Löfman, I., Szummer, K., Dahlström, U., Jernberg, T. & Lund, L. H. Associations with and prognostic impact of chronic kidney disease in heart failure with preserved, mid-range, and reduced ejection fraction. *Eur. J. Heart Fail.* **19**, 1606–1614 (2017).
5. Gheorghade, M. et al. Pathophysiologic targets in the early phase of acute heart failure syndromes. *Am. J. Cardiol.* **96**, 11G–17G (2005).
6. Schefold, J. C., Filippatos, G., Hasenfuss, G., Anker, S. D. & von Haehling, S. Heart failure and kidney dysfunction: epidemiology, mechanisms and management. *Nat. Rev. Nephrol.* **12**, 610–623 (2016).
7. Mullens, W. et al. Evaluation of kidney function throughout the heart failure trajectory—a position statement from the Heart Failure Association of the European Society of Cardiology. *Eur. J. Heart Fail.* **22**, 584–603 (2020).
8. McDonagh, T. A. et al. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. *Eur. Heart J.* **42**, 3599–3726 (2021).
9. Heidenreich, P. A. et al. 2022 AHA/ACC/HFSA guideline for the management of heart failure: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Circulation* **145**, e895–e1032 (2022).
10. CONSENSUS Trial Study Group Effects of enalapril on mortality in severe congestive heart failure. Results of the Cooperative North Scandinavian Enalapril Survival Study (CONSENSUS). *N. Engl. J. Med.* **316**, 1429–1435 (1987).
11. Yusuf, S. et al. Effect of enalapril on survival in patients with reduced left ventricular ejection fractions and congestive heart failure. *N. Engl. J. Med.* **325**, 293–302 (1991).
12. Yusuf, S. et al. Effect of enalapril on mortality and the development of heart failure in asymptomatic patients with reduced left ventricular ejection fractions. *N. Engl. J. Med.* **327**, 685–691 (1992).
13. Pfeffer, M. A. et al. Effects of candesartan on mortality and morbidity in patients with chronic heart failure: the CHARM-Overall programme. *Lancet* **362**, 759–766 (2003).
14. Pitt, B. et al. The effect of spironolactone on morbidity and mortality in patients with severe heart failure. *N. Engl. J. Med.* **341**, 709–717 (1999).
15. Zannad, F. et al. Eplerenone in patients with systolic heart failure and mild symptoms. *N. Engl. J. Med.* **364**, 11–21 (2011).
16. McMurray, J. J. et al. Angiotensin-neprilysin inhibition versus enalapril in heart failure. *N. Engl. J. Med.* **371**, 993–1004 (2014).
17. Solomon, S. D. et al. Angiotensin-neprilysin inhibition in heart failure with preserved ejection fraction. *N. Engl. J. Med.* **381**, 1609–1620 (2019).
18. Mehra, M. R. et al. The 2016 International Society for Heart Lung Transplantation listing criteria for heart transplantation: a 10-year update. *J. Heart Lung Transpl.* **35**, 1–23 (2016).
19. Damman, K. et al. Renal effects and associated outcomes during angiotensin-neprilysin inhibition in heart failure. *JACC Heart Fail.* **6**, 489–498 (2018).
20. Anker, S. D. et al. Empagliflozin in heart failure with a preserved ejection fraction. *N. Engl. J. Med.* **385**, 1451–1461 (2021).
21. Jhund, P. S. et al. Efficacy of dapagliflozin on renal function and outcomes in patients with heart failure with reduced ejection fraction: results of DAPA-HF. *Circulation* **143**, 298–309 (2021).
22. Mc Causland, F. R. et al. Dapagliflozin and kidney outcomes in patients with heart failure with mildly reduced or preserved ejection fraction: a prespecified analysis of the DELIVER randomized clinical trial. *JAMA Cardiol.* **8**, 56–65 (2023).
23. Sharma, A. et al. Cardiac and kidney benefits of empagliflozin in heart failure across the spectrum of kidney function: insights from the EMPEROR-Preserved trial. *Eur. J. Heart Fail.* **25**, 1337–1348 (2023).
24. Inker, L. A. et al. GFR slope as a surrogate end point for kidney disease progression in clinical trials: a meta-analysis of treatment effects of randomized controlled trials. *J. Am. Soc. Nephrol.* **30**, 1735–1745 (2019).
25. Lambers Heerspink, H. J. et al. Estimated GFR decline as a surrogate end point for kidney failure: a post hoc analysis from the Reduction of End Points in Non-Insulin-Dependent Diabetes With the Angiotensin II Antagonist Losartan (RENAAL) study and Irbesartan Diabetic Nephropathy Trial (IDNT). *Am. J. Kidney Dis.* **63**, 244–250 (2014).
26. Inker, L. A. et al. GFR decline as an alternative end point to kidney failure in clinical trials: a meta-analysis of treatment effects from 37 randomized trials. *Am. J. Kidney Dis.* **64**, 848–859 (2014).
27. Pocock, S. J., Ariti, C. A., Collier, T. J. & Wang, D. The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. *Eur. Heart J.* **33**, 176–182 (2012).
28. Redfors, B. et al. The win ratio approach for composite endpoints: practical guidance based on previous experience. *Eur. Heart J.* **41**, 4391–4399 (2020).
29. Little, D. J. et al. Validity and utility of a hierarchical composite end point for clinical trials of kidney disease progression: a review. *J. Am. Soc. Nephrol.* **34**, 1928–1935 (2023).
30. Heerspink, H. J. L. et al. Development and validation of a new hierarchical composite end point for clinical trials of kidney disease progression. *J. Am. Soc. Nephrol.* **34**, 2025–2038 (2023).
31. McMurray, J. J. V. et al. Dapagliflozin in patients with heart failure and reduced ejection fraction. *N. Engl. J. Med.* **381**, 1995–2008 (2019).
32. Solomon, S. D. et al. Dapagliflozin in heart failure with mildly reduced or preserved ejection fraction. *N. Engl. J. Med.* **387**, 1089–1098 (2022).
33. Packer, M. et al. Influence of endpoint definitions on the effect of empagliflozin on major renal outcomes in the EMPEROR-Preserved trial. *Eur. J. Heart Fail.* **23**, 1798–1799 (2021).
34. Adamson, C. et al. Initial decline (dip) in estimated glomerular filtration rate after initiation of dapagliflozin in patients with heart failure and reduced ejection fraction: insights from DAPA-HF. *Circulation* **146**, 438–449 (2022).
35. Chatur, S. et al. Variation in renal function following transition to sacubitril/valsartan in patients with heart failure. *J. Am. Coll. Cardiol.* **81**, 1443–1455 (2023).
36. Packer, M. Pitfalls in using estimated glomerular filtration rate slope as a surrogate for the effect of drugs on the risk of serious adverse renal outcomes in clinical trials of patients with heart failure. *Circ. Heart Fail.* **14**, e008537 (2021).
37. Inker, L. A. et al. A meta-analysis of GFR slope as a surrogate endpoint for kidney failure. *Nat. Med.* **29**, 1867–1876 (2023).
38. Vonesh, E. et al. Mixed-effects models for slope-based endpoints in clinical trials of chronic kidney disease. *Stat. Med.* **38**, 4218–4239 (2019).
39. McMurray, J. J. V. et al. A trial to evaluate the effect of the sodium-glucose co-transporter 2 inhibitor dapagliflozin on morbidity and mortality in patients with heart failure and reduced left ventricular ejection fraction (DAPA-HF). *Eur. J. Heart Fail.* **21**, 665–675 (2019).
40. McMurray, J. J. V. et al. The Dapagliflozin And Prevention of Adverse Outcomes in Heart Failure (DAPA-HF) trial: baseline characteristics. *Eur. J. Heart Fail.* **21**, 1402–1411 (2019).
41. Solomon, S. D. et al. Dapagliflozin in heart failure with preserved and mildly reduced ejection fraction: rationale and design of the DELIVER trial. *Eur. J. Heart Fail.* **23**, 1217–1225 (2021).
42. Solomon, S. D. et al. Baseline characteristics of patients with HF with mildly reduced and preserved ejection fraction: DELIVER trial. *JACC Heart Fail.* **10**, 184–197 (2022).
43. Ajufo, E., Nayak, A. & Mehra, M. R. Fallacies of using the win ratio in cardiovascular trials: challenges and solutions. *JACC Basic Transl. Sci.* **8**, 720–727 (2023).

44. Pocock, S. J. & Collier, T. J. Statistical appraisal of 6 recent clinical trials in cardiology: JACC state-of-the-art review. *J. Am. Coll. Cardiol.* **73**, 2740–2755 (2019).
45. Maurer, M. S. et al. Tafamidis treatment for patients with transthyretin amyloid cardiomyopathy. *N. Engl. J. Med.* **379**, 1007–1016 (2018).
46. Mack, M. J. et al. Transcatheter aortic-valve replacement with a balloon-expandable valve in low-risk patients. *N. Engl. J. Med.* **380**, 1695–1705 (2019).
47. Lopes, R. D. et al. Therapeutic versus prophylactic anticoagulation for patients admitted to hospital with COVID-19 and elevated D-dimer concentration (ACTION): an open-label, multicentre, randomised, controlled trial. *Lancet* **397**, 2253–2263 (2021).
48. Voors, A. A. et al. The SGLT2 inhibitor empagliflozin in patients hospitalized for acute heart failure: a multinational randomized trial. *Nat. Med.* **28**, 568–574 (2022).
49. Shah, S. J. et al. Atrial shunt device for heart failure with preserved and mildly reduced ejection fraction (REDUCE LAP-HF II): a randomised, multicentre, blinded, sham-controlled trial. *Lancet* **399**, 1130–1140 (2022).
50. Sorajja, P. et al. Transcatheter repair for patients with tricuspid regurgitation. *N. Engl. J. Med.* **388**, 1833–1842 (2023).
51. James, S. et al. Dapagliflozin in myocardial infarction without diabetes or heart failure. *NEJM Evid.* **3**, EVIDoa2300286 (2024).
52. Mentz, R. J. et al. Ferric carboxymaltose in heart failure with iron deficiency. *N. Engl. J. Med.* **389**, 975–986 (2023).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024



## Methods

### Study participants

In this post hoc study, we analyzed the DAPA-HF and DELIVER trials<sup>31,32</sup>. These were randomized, double-blind, placebo-controlled trials, and the trial designs and primary results have been published elsewhere<sup>31,32,39–42</sup>.

Briefly, DAPA-HF and DELIVER compared dapagliflozin to placebo in patients with a diagnosis of heart failure. Both trials enrolled patients with NYHA functional classes II–IV and elevated natriuretic peptide levels. The main difference between the two trials was that patients with an LVEF of  $\leq 40\%$  were randomized in the DAPA-HF trial and those with an LVEF  $> 40\%$  were randomized in the DELIVER trial. (DELIVER had evidence of structural heart disease, defined as either left atrial enlargement or left ventricular hypertrophy.) Key exclusion criteria included an eGFR lower than  $< 30 \text{ ml min}^{-1} 1.73 \text{ m}^{-2}$  in DAPA-HF and an eGFR  $< 25 \text{ ml min}^{-1} 1.73 \text{ m}^{-2}$  in DELIVER. In both trials, participants were randomized to receive dapagliflozin 10 mg once daily or a matching placebo. The median follow-up period was 1.5 years in the DAPA-HF trial and 2.3 years in the DELIVER trial.

Both trials were approved by the ethics committees at each investigative site and written informed consent was obtained from each participant.

### Study outcomes

The primary outcome was a composite of death from cardiovascular causes or worsening heart failure in DAPA-HF and DELIVER. In both trials, all-cause mortality was included as a secondary outcome, and a composite kidney outcome was included as a secondary outcome or prespecified exploratory outcome. All death events were adjudicated. The definition of ESKD in each trial was prespecified as a sustained eGFR  $< 15 \text{ ml min}^{-1} 1.73 \text{ m}^{-2}$ , chronic dialysis treatment or kidney transplantation in DAPA-HF and adverse event reporting, or a sustained eGFR  $< 15 \text{ ml min}^{-1} 1.73 \text{ m}^{-2}$  in DELIVER. The endpoints driven by the eGFR were derived from central laboratory results.

In this post hoc analysis, we examined a hierarchical composite outcome including the following components: all-cause mortality (tier 1); ESKD or eGFR  $< 15 \text{ ml min}^{-1} 1.73 \text{ m}^{-2}$  (tier 2); a decline in eGFR of  $\geq 57\%$  (tier 3); a decline in eGFR of  $\geq 50\%$  (tier 4); a decline in eGFR of  $\geq 40\%$  (tier 5); and participant-level eGFR slope (tier 6) (Extended Data Table 1). All-cause mortality was used for tier 1 in the hierarchy because of its ultimate clinical importance and its competing risk for the remaining outcomes. Considering the outcomes proposed by the international consensus definition of clinical trial outcomes for kidney disease, ESKD (or equivalent status) and decline in eGFR with different cutoffs were applied as tiers 2–5 (ref. 53). Decline in eGFR was applied as tier 6 because this has also been used for regulatory approval of treatment in some chronic kidney disease settings<sup>24–26</sup>. To address concerns regarding the lack of short-term verification of a change in eGFR due to the long interval between the scheduled study visits (and because some cutoffs were not verified as they were prespecified), declines in eGFR not requiring evidence that they were sustained eGFR were also evaluated. That is, change in eGFR (tiers 2–5) was evaluated as the time to the first meeting of the eGFR criterion based on the scheduled study visits, with the last laboratory assessment date used for censoring. eGFR was scheduled to be obtained at randomization, 14 days, 2 months, 4 months, 8 months, 12 months, 16 months, 20 months and 24 months in the DAPA-HF trial; and at randomization, 1 month, 4 months, 12 months, 24 months and 36 months in the DELIVER trial. The eGFR at randomization was used as the baseline eGFR to evaluate the change in eGFR; participants without baseline eGFR were excluded, that is, two participants in the DAPA-HF trial and one participant in the DELIVER trial. In this study, the original definition of ESKD in each study was used, alongside the aforementioned evaluation of change in eGFR.

As sensitivity analyses, we analyzed three additional models: sensitivity Model 1, excluding the component of a decline in eGFR of

$\geq 40\%$ , to evaluate outcomes less affected by the initial dip in eGFR due to the direct pharmacological action of dapagliflozin; sensitivity Model 2, excluding the component of a decline in eGFR of  $\geq 40\%$  and an eGFR slope to address additional concerns about the clinical relevance of the eGFR slope; and sensitivity Model 3, excluding all-cause mortality, which is more specific to kidney disease.

### Statistical analyses

To evaluate the effect of dapagliflozin across the range of LVEF, data were analyzed for the pooled dataset of DAPA-HF and DELIVER, and for each trial dataset separately.

Baseline characteristics were summarized according to the randomized group as the mean with s.d., or the median with the interquartile range for continuous variables and count with percentages for categorical variables. Continuous variables were compared using a *t*-test or Wilcoxon rank-sum test; categorical variables were compared using a chi-squared test. To determine the slope of change in eGFR for each individual patient over time according to the assigned treatment, two-slope, mixed-effect models accounting for the acute and chronic phases were applied using the eGFR data obtained at all scheduled visits<sup>30</sup>. The acute phase was defined as the period up to the first postrandomization visit (14 days in DAPA-HF and 1 month in DELIVER) when the acute treatment effect on the eGFR was considered fully present. These models were adjusted for baseline eGFR values, randomized treatment, visit time, diabetes status, spline variable corresponding to the days since the acute phase, the interaction of treatment and visit time, and the interaction of treatment and spline, without an intercept term. The distributions of the individual eGFR slopes were drawn using violin plots.

The unmatched win statistics method, in which every patients in the dapagliflozin group was paired and compared with every patient in the placebo group, was used<sup>27</sup>; pairs representing the product of the number of individuals in the dapagliflozin group and placebo group were created and compared. Comparisons were made in ascending order of event tier (from 1 to 6); once a tier was settled, the next tier was not assessed; if the last tier was not settled, the comparison pair was considered a tie (Extended Data Fig. 7). In tiers 1–5, the time to first event was compared during a fixed follow-up period; censoring earlier than the defined fixed follow-up period was considered censoring at the fixed follow-up period to address the effect of censoring distributions on win statistics results<sup>27,54–58</sup>. Fixed follow-up periods were defined as 720 days in DAPA-HF and 1,080 days in DELIVER, considering the scheduled visits and follow-up period. In tier 6, the participant-level eGFR slope, which was calculated using data within these fixed follow-up periods, was compared as a continuous variable in each pair (that is, the patient with a shallower eGFR slope is the winner); thus, in the model including the eGFR slope, tied pairs did not exist. The proportions of win pairs ( $P_W$ ), loss pairs ( $P_L$ ) and tied pairs ( $P_T$ ) for participants assigned to dapagliflozin were obtained;  $P_W$  is the number of win pairs divided by the total number of pairs  $n_D \times n_P$  where  $n_D$  and  $n_P$  are the sample sizes in the dapagliflozin and placebo group, similarly for  $P_L$  and  $P_T$ . The method outlined by Pocock et al.<sup>27</sup> and the corresponding variances based on the U-statistic-based method by Dong et al.<sup>59</sup> were used to compute the win ratio. Because of a shortcoming of the win ratio that ignores ties when comparing pairs to obtain the win ratio, we calculated the ‘win odds’ for sensitivity Model 2, which is a modification of the win ratio accounting for ties<sup>60,61</sup>. Net benefit was also reported, which is the difference between the proportion of win and loss pairs<sup>58</sup>. We calculated four win statistics (win ratio, net benefit, win odds and win probability) defined as: win ratio,  $P_W/P_L$ ; net benefit,  $P_W - P_L$ ; win odds,  $(P_W + 0.5 P_T)/(P_L + 0.5 P_T)$ ; and win probability,  $P_W + 0.5 P_T$ . Thus, in the main model, sensitivity Model 1 and sensitivity Model 3, where tied pairs do not exist, the win ratio is identical to the win odds. A win ratio represents the ratio of the proportion of win pairs to the proportion of loss pairs; a win rate greater than 1 with a lower 95% CI greater than 1

indicates that dapagliflozin is better than placebo. Because the win or loss proportion depends on the duration of follow-up and the censoring distribution, we plotted these trends over time every 10 days<sup>55,62</sup>. This plot was drawn only for tiers 1–5 because the eGFR slope was calculated based on data across the fixed follow-up period, meaning it was not possible to report an eGFR slope at a specific time point and illustrate the proportion of the wins or losses over time for this component of the composite outcome.

We also evaluated the component of the kidney hierarchical composite outcome up to the aforementioned fixed follow-up period using conventional statistical approaches to compare these results with the ones from the win statistic. Cox proportional hazards models were used to compute the HRs (to aid direct comparison, these are presented as 1/HR) and Kaplan–Meier curves were plotted.

Consistent with the prespecified stratification variables in each respective trial, win statistics and Cox proportional hazards models were stratified according to diabetes status and trial in the pooled dataset<sup>31,32</sup>.

The sample size requirements and statistical power of the hierarchical composite endpoint (main model) were compared using bootstrap resampling of the pooled dataset with the time-to-first composite endpoint (all-cause mortality, ESKD or eGFR <15 ml min<sup>-1</sup> 1.73 m<sup>-2</sup>, or decline in eGFR of ≥40%) and eGFR slope to detect the observed treatment effect for each endpoint. The resampling procedure was performed with 1,000 iterations at each sample size ( $n = 200, 500$  and increments of 500 until 3,000).

All analyses were conducted using STATA v.17.0 and R v.4.2.2.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

AstraZeneca's data-sharing policy is described at <https://astrazeneca-grouptrials.pharmacm.com/ST/Submission/Disclosure>. Researchers need to submit a request to access anonymized patient-level clinical data, aggregated clinical data or anonymized clinical study documents through Vivli's web-based data request platform (<https://vivli.org/>). An independent scientific review board will review requests. Timelines vary per request and can take up to a year upon full submission of the request for analysis, decision, anonymization and sharing of the requested data or documents.

### Code availability

The key code to obtain the eGFR slope was published by Heerspink et al.<sup>30</sup>. Win statistics were conducted using the WINS package of R (<https://cran.r-project.org/web/packages/WINS/index.html>). The detailed code used to generate the findings of the present study is available from the corresponding author (john.mcmurray@glasgow.ac.uk) upon request from qualified researchers in this field. Researchers are asked to provide information on their affiliation and experience in this field and how they intend to use the code. The timelines vary per request and can take up to 6 months upon submission of the request.

### References

- Levin, A. et al. International consensus definitions of clinical trial outcomes for kidney failure: 2020. *Kidney Int.* **98**, 849–859 (2020).
- Rauch, G., Jahn-Eimermacher, A., Brannath, W. & Kieser, M. Opportunities and challenges of combined effect measures based on prioritized outcomes. *Stat. Med.* **33**, 1104–1120 (2014).
- Oakes, D. On the win-ratio statistic in clinical trials with multiple types of event. *Biometrika* **103**, 742–745 (2016).
- Bebu, I. & Lachin, J. M. Large sample inference for a win ratio analysis of a composite outcome based on prioritized components. *Biostatistics* **17**, 178–187 (2016).
- Péron, J., Buyse, M., Ozenne, B., Roche, L. & Roy, P. An extension of generalized pairwise comparisons for prioritized outcomes in the presence of censoring. *Stat. Methods Med. Res.* **27**, 1230–1239 (2018).
- Dong, G. et al. Win statistics (win ratio, win odds, and net benefit) can complement one another to show the strength of the treatment effect on time-to-event outcomes. *Pharm. Stat.* **22**, 20–33 (2023).
- Dong, G., Li, D., Ballerstedt, S. & Vandemeulebroecke, M. A generalized analytic solution to the win ratio to analyze a composite endpoint considering the clinical importance order among components. *Pharm. Stat.* **15**, 430–437 (2016).
- Dong, G. et al. The win ratio: on interpretation and handling of ties. *Stat. Biopharm. Res.* **12**, 99–106 (2020).
- Brunner, E., Vandemeulebroecke, M. & Mütze, T. Win odds: an adaptation of the win ratio to include ties. *Stat. Med.* **40**, 3367–3384 (2021).
- Finkelstein, D. M. & Schoenfeld, D. A. Graphing the Win Ratio and its components over time. *Stat. Med.* **38**, 53–61 (2019).

### Acknowledgements

P.S.J. and J.J.V.M. are supported by a British Heart Foundation Centre of Research Excellence grant no. RE/18/6/34217 and the Vera Melrose Heart Failure Research Fund. T.K. is supported by grant no. 20K17112 from Grant-in-Aid for Scientific Research. DAPA-HF and DELIVER were funded by AstraZeneca; however, the analyses and writing of the manuscript were conducted independently at the University of Glasgow.

### Author contributions

T.K., P.S.J. and J.J.V.M. conceived and designed the study. T.K., P.S.J. and S.B.G. performed the data analyses. T.K. wrote the first draft of the paper. J.J.V.M. and S.D.S. oversaw and supervised the study. All authors contributed to the interpretation of the data, provided critical feedback on the paper drafts and approved the manuscript for submission.

### Competing interests

T.K. has received speaker fees from Abbott, Ono Pharma, Otsuka Pharma, Novartis, AstraZeneca, Bristol Myers Squibb, Boehringer Ingelheim and Abiomed. P.S.J. reports speakers' fees from AstraZeneca, Novartis, Alkem Metabolics, ProAdWise Communications, Sun Pharmaceuticals and Intas Pharmaceuticals; advisory board fees from AstraZeneca, Boehringer Ingelheim and Novartis; and research funding from AstraZeneca, Boehringer Ingelheim and Analog Devices Inc; P.S.J.'s employer, the University of Glasgow, has been remunerated for clinical trial work by AstraZeneca, Bayer, Novartis and Novo Nordisk. He is a director of Global Clinical Trial Partners. M.Y. reports travel grants from AstraZeneca. S.B.G. is an employee and shareholder of AstraZeneca. B.L.C. has received consulting fees from Boehringer Ingelheim. F.R.M. reports research funding from the National Institute of Diabetes and Digestive and Kidney Diseases, Satellite Healthcare, Fifth Eye, Novartis and Lexicon, which is paid directly to his institution; he reports consulting fees from GSK and Zydus Therapeutics. M.V. has received research grant support, served on advisory boards or had speaker engagements with American Regent, Amgen, AstraZeneca, Bayer, Baxter Healthcare, Boehringer Ingelheim, Chiesi, Cytokinetics, Lexicon Pharmaceuticals, Merck, Novartis, Novo Nordisk, Pharmacosmos, Relypsa, Roche Diagnostics, Sanofi and Tricog Health; he sits on clinical trial committees for studies sponsored by AstraZeneca, Galmed, Novartis, Bayer, Occlutech and Impulse Dynamics. H.J.L.H. has served as a consultant for AbbVie, Astellas, AstraZeneca, Boehringer Ingelheim, Fresenius, Gilead, Janssen, Merck and Mitsubishi Tanabe; he has received grant support from AbbVie, AstraZeneca, Boehringer Ingelheim and Janssen. S.D.S. has received research grants from

Actelion, Alnylam, Amgen, AstraZeneca, Bellerophon, Bayer, Bristol Myers Squibb, Celladon, Cytokinetics, Eidos, Gilead, GSK, Ionis, Lilly, Mesoblast, MyoKardia, National Institutes of Health/National Heart, Lung, and Blood Institute, Neurotronik, Novartis, Novo Nordisk, Respicardia, Sanofi Pasteur, Theracos and Us2.ai; he has consulted for Abbott, Action, Akros, Alnylam, Amgen, Arena, AstraZeneca, Bayer, Boehringer Ingelheim, Bristol Myers Squibb, Cardior, Cardurion, Corvia, Cytokinetics, Daiichi-Sankyo, GSK, Lilly, Merck, Myokardia, Novartis, Roche, Theracos, Quantum Genomics, Cardurion, Janssen, Cardiac Dimensions, Tenaya, Sanofi Pasteur, Dinaqor, Trembeau, CellProthera, Moderna, American Regent and Sarepta. J.J.V.M. reports payments through Glasgow University from work on clinical trials, consulting and other activities from Amgen, AstraZeneca, Bayer, Cardurion, Cytokinetics, GSK, KBP Biosciences and Novartis. He reports personal consultancy fees from Alnylam Pharmaceuticals, Bayer, Bristol Myers Squibb, George Clinical PTY Ltd, Ionis Pharmaceuticals, Novartis, Regeneron Pharmaceuticals and the River 2 Renal Corporation. He receives personal lecture fees from Abbott, Alkem Metabolics, AstraZeneca, Blue Ocean Scientific Solutions Ltd, Boehringer Ingelheim, Canadian Medical and Surgical Knowledge, Emcure Pharmaceuticals Ltd, Eris Lifesciences, European Academy of CME, Hikma Pharmaceuticals, Imagica Health, Intas Pharma, J.B. Chemicals & Pharmaceuticals Ltd, Lupin Pharma, Medscape/

TheHeart.Org, ProAdWise Communications, Radcliffe Cardiology, Sun Pharmaceuticals, The Corpus, the Translation Research Group and the Translational Medicine Academy. He is a director of Global Clinical Trial Partners.

### Additional information

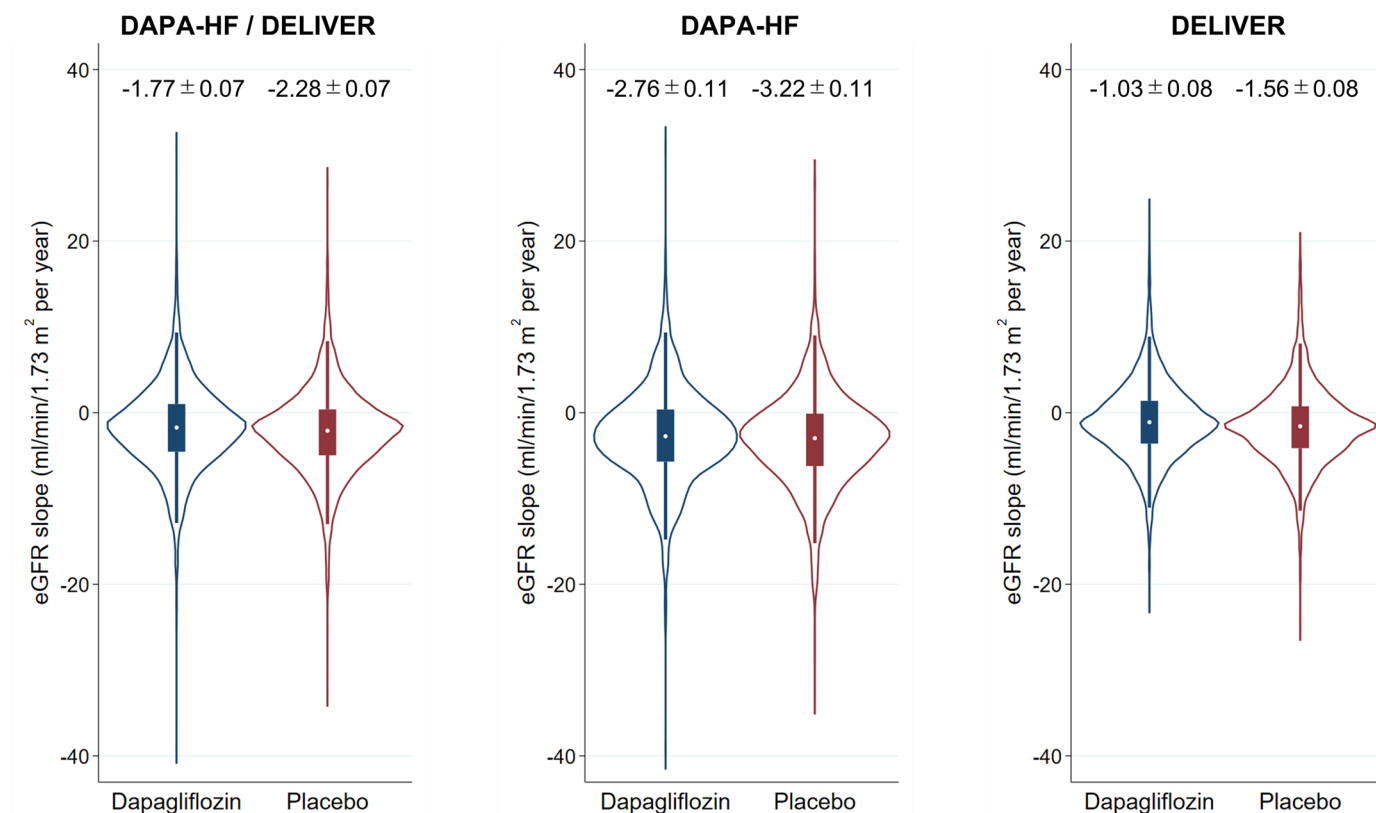
**Extended data** is available for this paper at <https://doi.org/10.1038/s41591-024-02941-8>.

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41591-024-02941-8>.

**Correspondence and requests for materials** should be addressed to John J. V. McMurray.

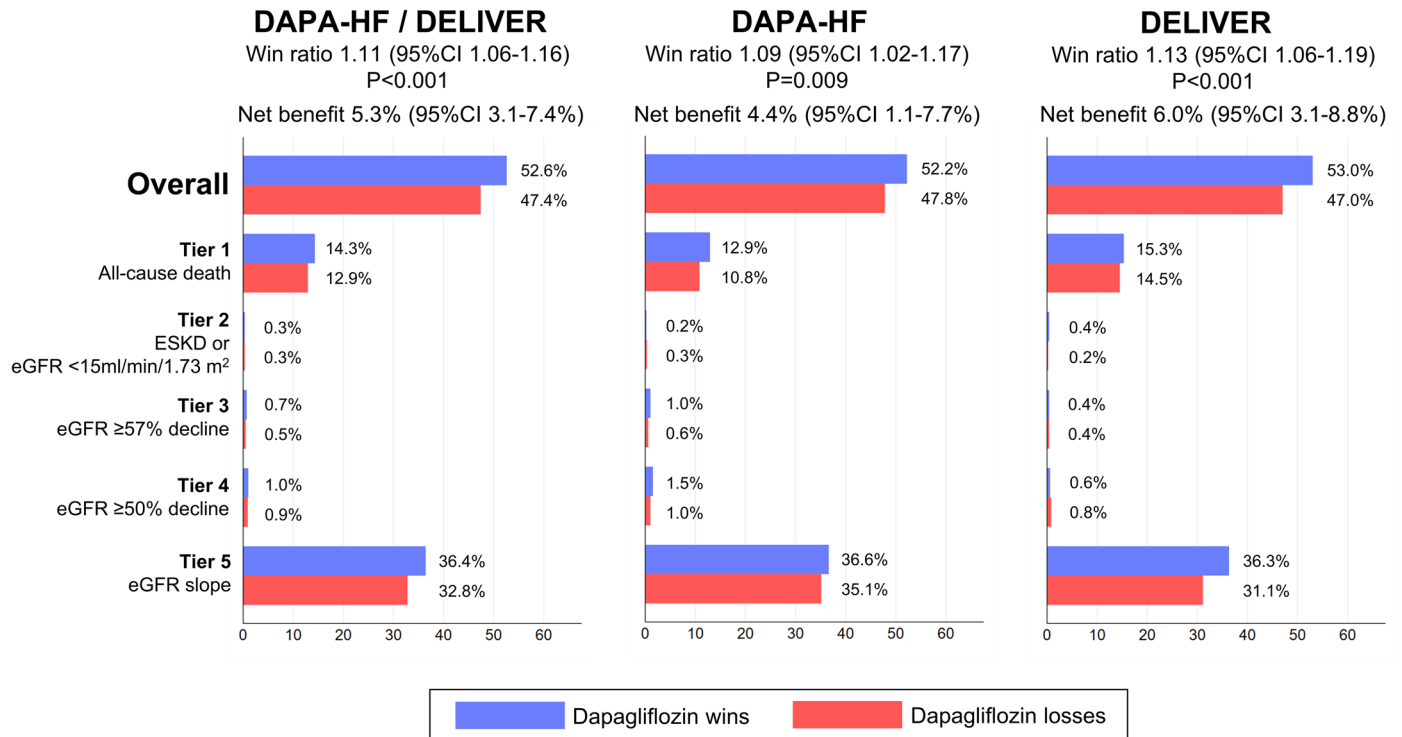
**Peer review information** *Nature Medicine* thanks Peter van der Meer and the other anonymous reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Ming Yang, in collaboration with the *Nature Medicine* team.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).



**Extended Data Fig. 1 | Distribution of the overall eGFR slope in each trial.** The numbers above the violin plots are the mean ± standard error of the eGFR slope. The white dots at the center of the violin plots show the median eGFR slope and the boxes around these indicate the 25th to 75th percentile values. The whiskers indicate the full range of the eGFR slope. The 'violin' shape is a density estimation showing the distribution of the eGFR slope values. Wider sections of the violin plot represent a higher probability and the narrower sections represent a lower probability of patients having the value. The eGFR slopes between randomized

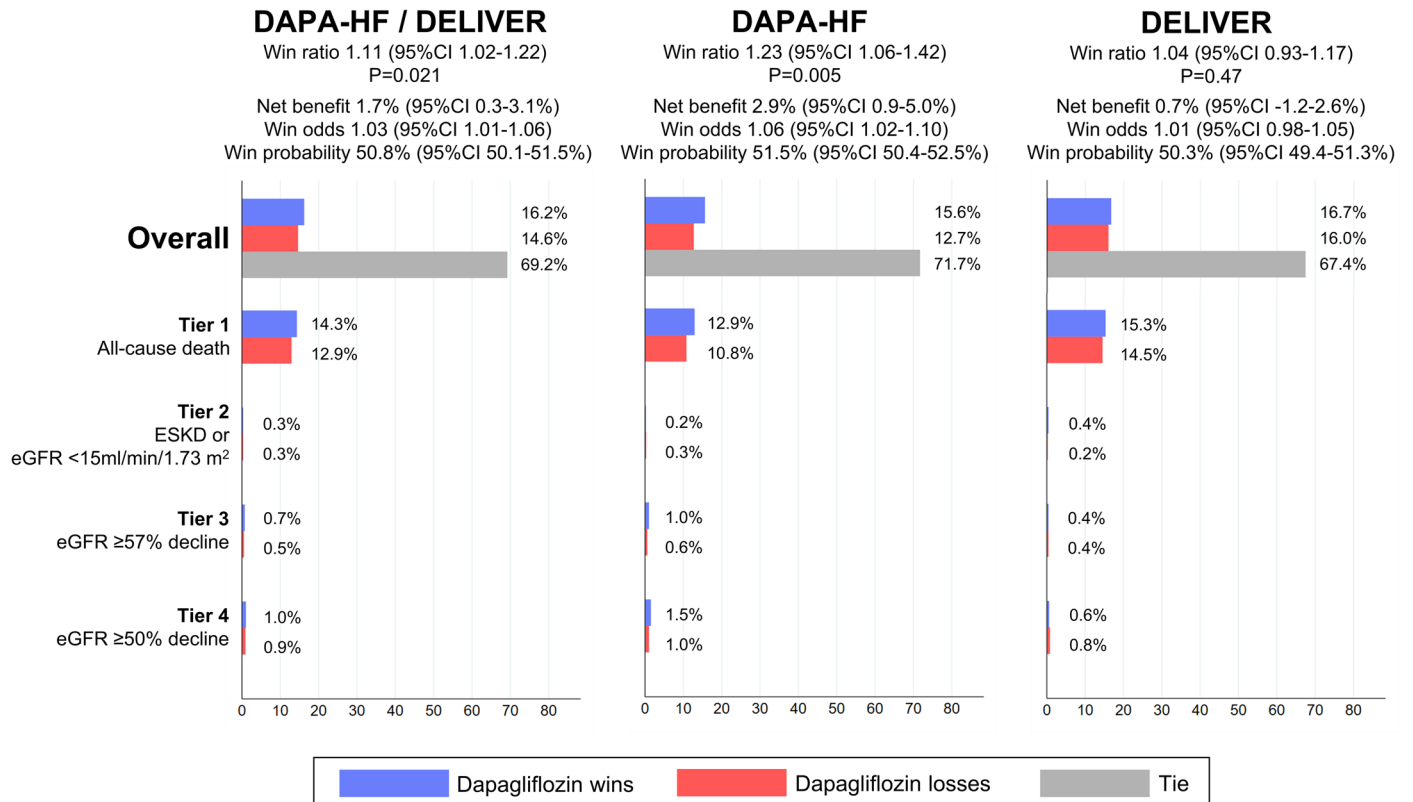
treatment groups are compared by the two-sided t-test. Adjustments are not made for multiple comparisons. The exact p-values were 0.0000002 in the pooled dataset, 0.004 in DAPA-HF, and 0.000004 in DELIVER. CI, confidence interval; DAPA-HF, Dapagliflozin and Prevention of Adverse Outcomes in Heart Failure; DELIVER, Dapagliflozin Evaluation to Improve the Lives of Patients with Preserved Ejection Fraction Heart Failure trial; eGFR, estimated glomerular filtration rate.



**Extended Data Fig. 2 | Effect of dapagliflozin on the hierarchical composite outcome in sensitivity model 1.** Win statistics are two-sided. Models are stratified by diabetes status (and by trial in the pooled dataset). Adjustments are not made for multiple comparisons. The exact p-values were 0.000002 in the pooled dataset and 0.00005 in DELIVER. CI, confidence interval; DAPA-HF,

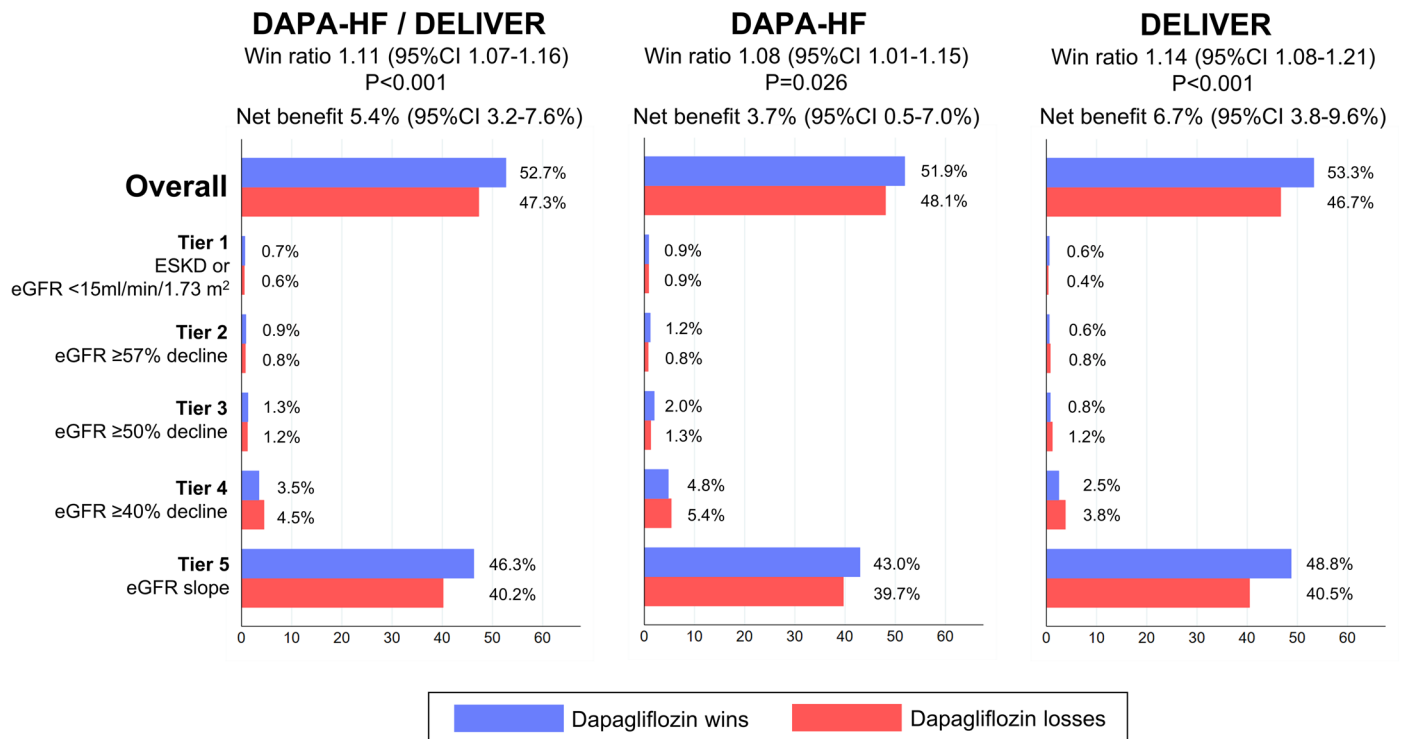
Dapagliflozin and Prevention of Adverse Outcomes in Heart Failure; DELIVER, Dapagliflozin Evaluation to Improve the Lives of Patients with Preserved Ejection Fraction Heart Failure trial; eGFR, estimated glomerular filtration rate; ESKD, end-stage kidney disease.





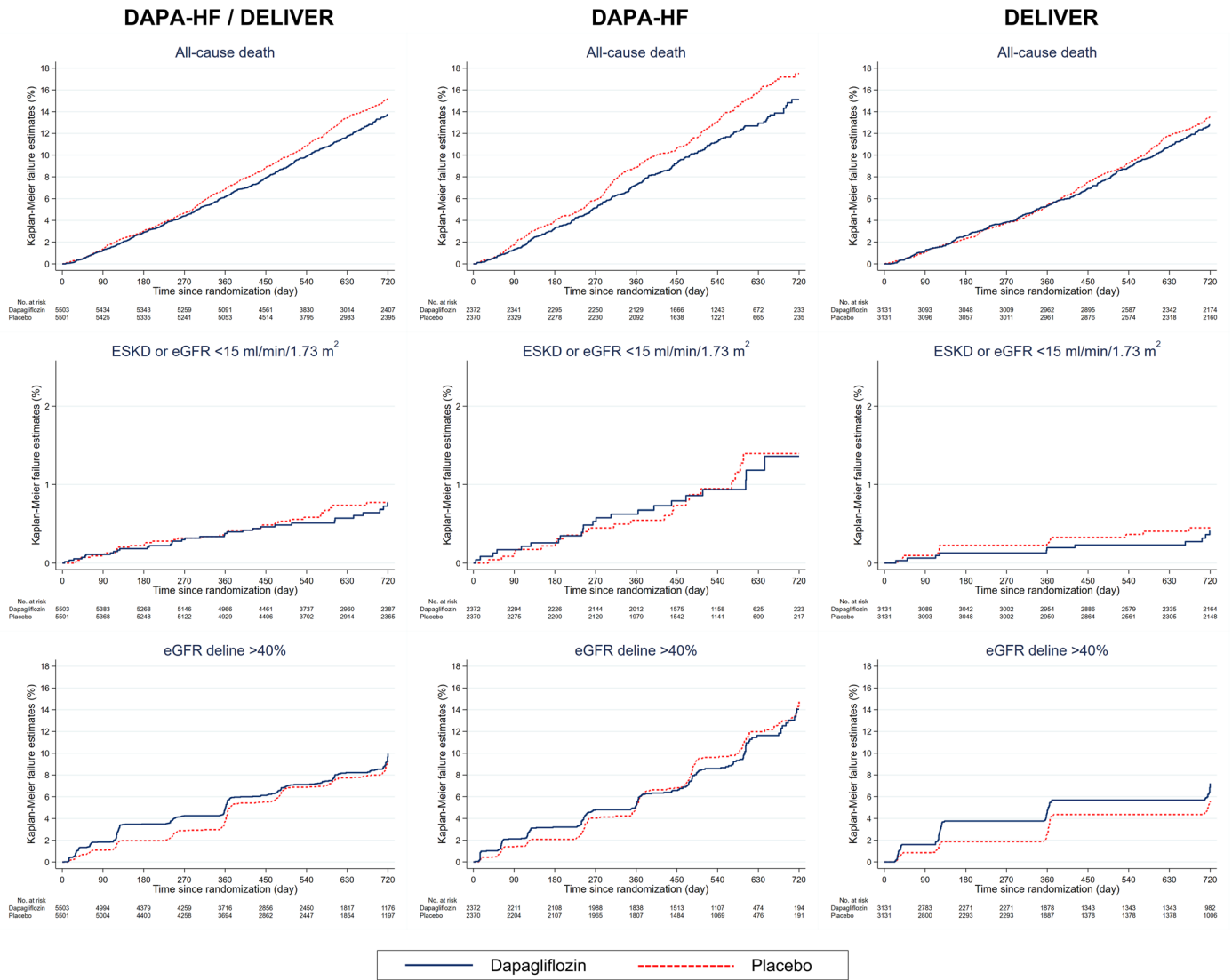
**Extended Data Fig. 3 | Effect of dapagliflozin on the hierarchical composite outcome in sensitivity model 2.** Win statistics are two-sided. Models are stratified by diabetes status (and by trial in the pooled dataset). Adjustments are not made for multiple comparisons. CI, confidence interval; DAPA-HF,

Dapagliflozin and Prevention of Adverse Outcomes in Heart Failure; DELIVER, Dapagliflozin Evaluation to Improve the Lives of Patients with Preserved Ejection Fraction Heart Failure trial; eGFR, estimated glomerular filtration rate; ESKD, end-stage kidney disease.

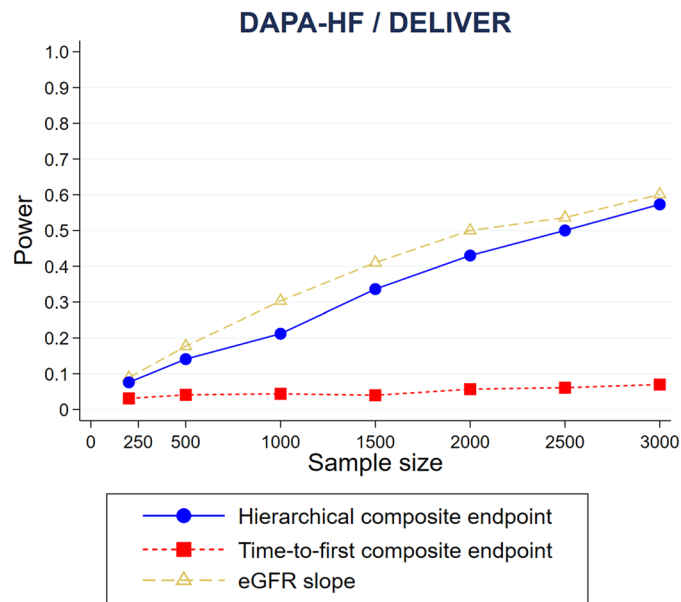


**Extended Data Fig. 4 | Effect of dapagliflozin on the hierarchical composite outcome in sensitivity model 3.** Win statistics are two-sided. Models are stratified by diabetes status (and by trial in the pooled dataset). Adjustments are not made for multiple comparisons. The exact p-values were 0.0000009 in the pooled dataset and 0.000005 in DELIVER. CI, confidence interval; DAPA-HF,

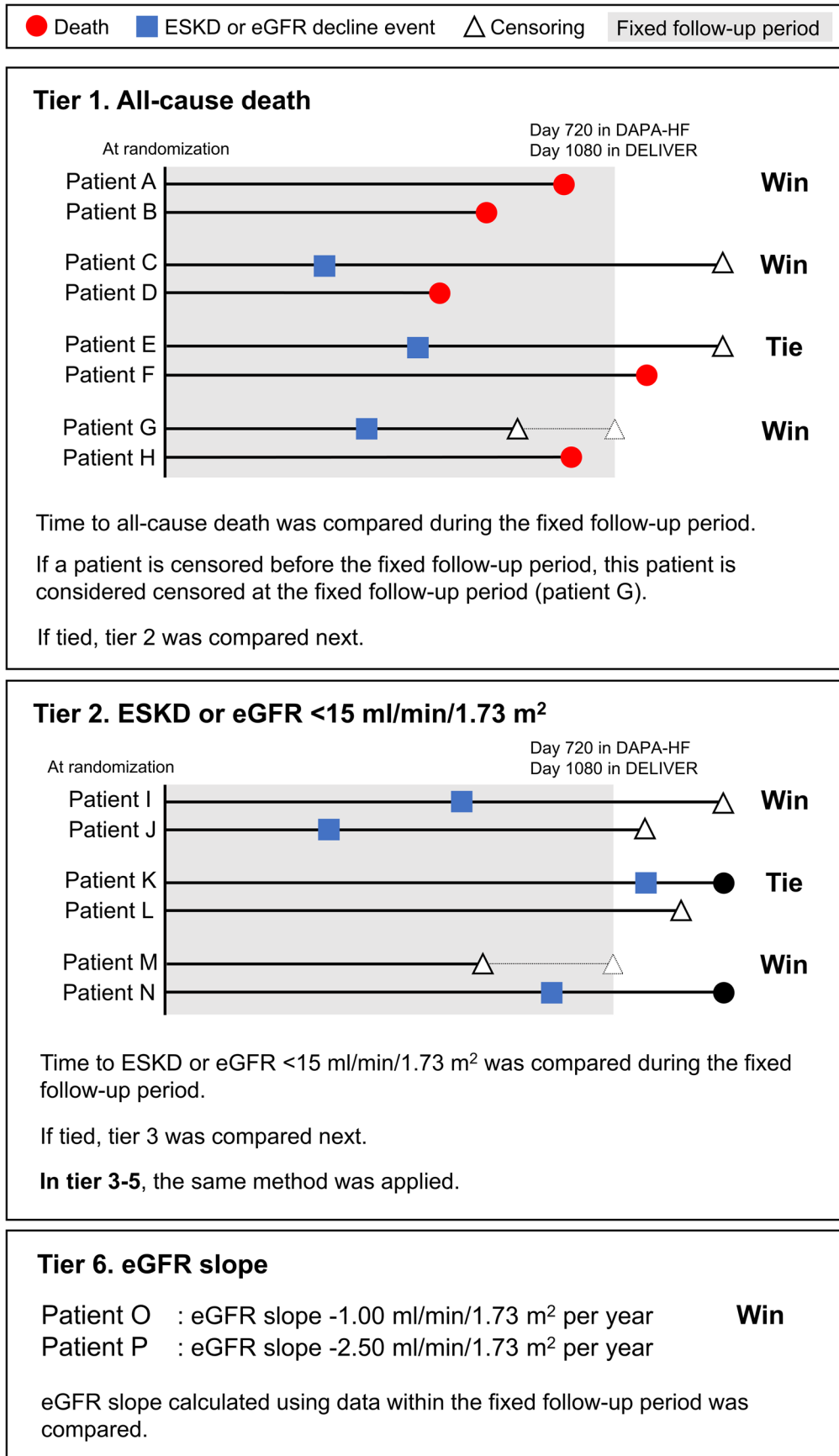
Dapagliflozin and Prevention of Adverse Outcomes in Heart Failure; DELIVER, Dapagliflozin Evaluation to Improve the Lives of Patients with Preserved Ejection Fraction Heart Failure trial; eGFR, estimated glomerular filtration rate; ESKD, end-stage kidney disease.



**Extended Data Fig. 5 | Effect of dapagliflozin in Kaplan-Meier plots.** DAPA-HF, Dapagliflozin and Prevention of Adverse Outcomes in Heart Failure; DELIVER, Dapagliflozin Evaluation to Improve the Lives of Patients with Preserved Ejection Fraction Heart Failure trial; eGFR, estimated glomerular filtration rate; ESKD, end-stage kidney disease.



**Extended Data Fig. 6 | Curves for sample size and statistical power in the pooled dataset of DAPA-HF and DELIVER.** DAPA-HF, Dapagliflozin and Prevention of Adverse Outcomes in Heart Failure; DELIVER, Dapagliflozin Evaluation to Improve the Lives of Patients with Preserved Ejection Fraction Heart Failure trial; eGFR, estimated glomerular filtration rate.



**Extended Data Fig. 7 | Testing the hierarchical composite outcome.** DAPA-HF, Dapagliflozin and Prevention of Adverse Outcomes in Heart Failure; DELIVER, Dapagliflozin Evaluation to Improve the Lives of Patients with Preserved Ejection Fraction Heart Failure trial; eGFR, estimated glomerular filtration rate; ESKD, end-stage kidney disease.



## Extended Data Table 1 | Summary of the hierarchical composite outcomes

Outcome Model	Tier	Component of outcomes
Main model	1	All-cause death
	2	ESKD or eGFR < 15ml/min/1.73 m <sup>2</sup>
	3	eGFR ≥57% decline
	4	eGFR ≥50% decline
	5	eGFR ≥40% decline
	6	eGFR slope
Sensitivity model 1	1	All-cause death
	2	ESKD or eGFR < 15ml/min/1.73 m <sup>2</sup>
	3	eGFR ≥57% decline
	4	eGFR ≥50% decline
	5	eGFR slope
Sensitivity model 2	1	All-cause death
	2	ESKD or eGFR < 15ml/min/1.73 m <sup>2</sup>
	3	eGFR ≥57% decline
	4	eGFR ≥50% decline
Sensitivity model 3	1	ESKD or eGFR < 15ml/min/1.73 m <sup>2</sup>
	2	eGFR ≥57% decline
	3	eGFR ≥50% decline
	4	eGFR ≥40% decline
	5	eGFR slope

eGFR, estimated glomerular filtration rate; ESKD, end-stage kidney disease.

Corresponding author(s): John JV McMurrayLast updated by author(s): Mar 5, 2024

## Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection

The key code to obtain the eGFR slope was published in the paper of Heerspink, H.J.L., et al. Development and Validation of a New Hierarchical Composite End Point for Clinical Trials of Kidney Disease Progression. *J Am Soc Nephrol* 34, 2025-2038 (2023). Win statistics were conducted using the WINS package of R (<https://cran.r-project.org/web/packages/WINS/index.html>). The detailed code used to generate the findings of the present study is available from the corresponding author (john.mcmurray@glasgow.ac.uk) on request from qualified researchers in this field. Researchers are asked to provide information on their affiliation and experience in this field and how they intend to use the code. The timelines vary per request and can take up to 6 months upon submission of the request.

Data analysis

See above.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

AstraZeneca's data-sharing policy is described at <https://astrazenecagrouptrials.pharmacm.com/ST/Submission/Disclosure>.

Researchers need to submit a request to access anonymized patient-level clinical data, aggregated clinical data, and/or anonymized clinical study documents through Vivli's web-based data request platform (<https://vivli.org/>).

An independent Scientific Review Board is available to review requests.

The timelines vary per request and can take up to a year upon full submission of the request for analysis, decision, anonymisation and sharing of the requested data or documents.

## Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

### Reporting on sex and gender

Biological sex was reported by investigators. Gender was not reported. Both sexes were enrolled (65% male and 35% female in this study). Thus, findings can be applied to both sexes. Analysis by sex was not performed because there is no difference in the effect of dapagliflozin by sex.

1: Wang X, Vaduganathan M, Claggett BL, Hegde SM, Pabon M, Kulac IJ, Vardeny O, O'Meara E, Zieroth S, Katova T, McGrath MM, Pouleur AC, Jhund PS, Desai AS, Inzucchi SE, Kosiborod MN, de Boer RA, Kober L, Sabatine MS, Martinez FA, Ponikowski P, Shah SJ, Hernandez AF, Langkilde AM, McMurray JJV, Solomon SD, Lam CSP. Sex Differences in Characteristics, Outcomes, and Treatment Response With Dapagliflozin Across the Range of Ejection Fraction in Patients With Heart Failure: Insights From DAPA-HF and DELIVER. *Circulation*. 2023 Feb 21;147(8):624-634. doi: 10.1161/CIRCULATIONAHA.122.062832. Epub 2022 Nov 7. PMID: 36342789; PMCID: PMC9974767.

2: Butt JH, Docherty KF, Petrie MC, Schou M, Kosiborod MN, O'Meara E, Katova T, Ljungman CEA, Diez M, Ogunniyi MO, Langkilde AM, Sjöstrand M, Lindholm D, Bengtsson O, Martinez FA, Ponikowski P, Sabatine MS, Solomon SD, Jhund PS, McMurray JJV, Køber L. Efficacy and Safety of Dapagliflozin in Men and Women With Heart Failure With Reduced Ejection Fraction: A Prespecified Analysis of the Dapagliflozin and Prevention of Adverse Outcomes in Heart Failure Trial. *JAMA Cardiol*. 2021 Jun 1;6(6):678-689. doi: 10.1001/jamacardio.2021.0379. PMID: 33787831; PMCID: PMC8014207.

### Reporting on race, ethnicity, or other socially relevant groupings

Race was classified according to the United States Food and Drug Administration's guidance. Overall, 71% of patients were classified as White, 22% Asian, 3% Black or African-American, and 4% Other. Investigators were asked to report race as self-identified by patients. Race was not taken into account in the present analyses.

### Population characteristics

HFrEF affects men more often than women (ratio of 2:1 or 3:1) whereas HFmrEF/HFpEF affects men and women similarly (or there is even a small female preponderance). The prevalence of both types of heart failure increases steeply with age but patients with HFmrEF/HFpEF are older, on average, than patients with HFrEF.

### Recruitment

Recruitment was conducted at 410 sites in 20 countries (DAPA-HF) and at 353 sites in 20 countries (DELIVER) according to the protocol inclusion and exclusion criteria.

### Ethics oversight

Both trials were registered at ClinicalTrials.gov before enrolment of the first patients, both protocols were approved by ethics committees at all sites, and all patients provided written informed consent, as described in the principal results manuscripts (see below):

1: McMurray JJV, Solomon SD, Inzucchi SE, Køber L, Kosiborod MN, Martinez FA, Ponikowski P, Sabatine MS, Anand IS, Böhlhávek J, Böhm M, Chiang CE, Chopra VK, de Boer RA, Desai AS, Diez M, Drozd J, Dukát A, Ge J, Howlett JG, Katova T, Kitakaze M, Ljungman CEA, Merkely B, Nicolau JC, O'Meara E, Petrie MC, Vinh PN, Schou M, Tereshchenko S, Verma S, Held C, DeMets DL, Docherty KF, Jhund PS, Bengtsson O, Sjöstrand M, Langkilde AM; DAPA-HF Trial Committees and Investigators. Dapagliflozin in Patients with Heart Failure and Reduced Ejection Fraction. *N Engl J Med*. 2019 Nov 21;381(21):1995-2008. doi: 10.1056/NEJMoa1911303.

2: Solomon SD, McMurray JJV, Claggett B, de Boer RA, DeMets D, Hernandez AF, Inzucchi SE, Kosiborod MN, Lam CSP, Martinez F, Shah SJ, Desai AS, Jhund PS, Belohlavek J, Chiang CE, Borleffs CJW, Comin-Colet J, Dobeanu D, Drozd J, Fang JC, Alcocer-Gamba MA, Al Habeeb W, Han Y, Cabrera Honorio JW, Janssens SP, Katova T, Kitakaze M, Merkely B, O'Meara E, Saraiva JFK, Tereshchenko SN, Thierer J, Vaduganathan M, Vardeny O, Verma S, Pham VN, Wilderäng U, Zaozerska N, Bachus E, Lindholm D, Petersson M, Langkilde AM; DELIVER Trial Committees and Investigators. Dapagliflozin in Heart Failure with

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences  Behavioural & social sciences  Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

Sample size was 11007, 4744 from DAPA-HF and 6263 from DELIVER. The sample size and power calculations are described in the following:

1: McMurray JJV, Solomon SD, Inzucchi SE, Køber L, Kosiborod MN, Martinez FA, Ponikowski P, Sabatine MS, Anand IS, Bělohávek J, Böhm M, Chiang CE, Chopra VK, de Boer RA, Desai AS, Díez M, Drozd J, Dukát A, Ge J, Howlett JG, Katova T, Kitakaze M, Ljungman CEA, Merkely B, Nicolau JC, O'Meara E, Petrie MC, Vinh PN, Schou M, Tereshchenko S, Verma S, Held C, DeMets DL, Docherty KF, Jhund PS, Bengtsson O, Sjöstrand M, Langkilde AM; DAPA-HF Trial Committees and Investigators. Dapagliflozin in Patients with Heart Failure and Reduced Ejection Fraction. N Engl J Med. 2019 Nov 21;381(21):1995-2008. doi: 10.1056/NEJMoa1911303.

2: Solomon SD, McMurray JJV, Claggett B, de Boer RA, DeMets D, Hernandez AF, Inzucchi SE, Kosiborod MN, Lam CSP, Martinez F, Shah SJ, Desai AS, Jhund PS, Belohlavek J, Chiang CE, Borleffs CJW, Comin-Colet J, Dobreanu D, Drozd J, Fang JC, Alcocer-Gamba MA, Al Habeeb W, Han Y, Cabrera Honorio JW, Janssens SP, Katova T, Kitakaze M, Merkely B, O'Meara E, Saraiva JFK, Tereshchenko SN, Thierer J, Vaduganathan M, Vardeny O, Verma S, Pham VN, Wilderäng U, Zaozerska N, Bachus E, Lindholm D, Petersson M, Langkilde AM; DELIVER Trial Committees and Investigators. Dapagliflozin in Heart Failure with Mildly Reduced or Preserved Ejection Fraction. N Engl J Med. 2022 Sep 22;387(12):1089-1098. doi: 10.1056/NEJMoa2206286. Epub 2022 Aug 27. PMID: 36027570.

Data exclusions

None.

Replication

Not reproducible because the data contains patient-specific information for over 11,000 cases

Randomization

Both were double blind placebo controlled prospective randomized trials.

Blinding

Both trials were double blind.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

- | n/a                                 | Involved in the study                                  |
|-------------------------------------|--|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Antibodies                    |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Eukaryotic cell lines         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Clinical data      |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Plants                        |

### Methods

- | n/a                                 | Involved in the study                           |
|-------------------------------------|---|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq               |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |

## Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration

NCT03036124 and NCT03619213

Study protocol

Statistical plans have been submitted with manuscript. See also,

1: McMurray JJV, Solomon SD, Inzucchi SE, Køber L, Kosiborod MN, Martinez FA, Ponikowski P, Sabatine MS, Anand IS, Bělohávek J,

Böhm M, Chiang CE, Chopra VK, de Boer RA, Desai AS, Diez M, Drozd J, Dukát A, Ge J, Howlett JG, Katova T, Kitakaze M, Ljungman CEA, Merkely B, Nicolau JC, O'Meara E, Petrie MC, Vinh PN, Schou M, Tereshchenko S, Verma S, Held C, DeMets DL, Docherty KF, Jhund PS, Bengtsson O, Sjöstrand M, Langkilde AM; DAPA-HF Trial Committees and Investigators. Dapagliflozin in Patients with Heart Failure and Reduced Ejection Fraction. *N Engl J Med*. 2019 Nov 21;381(21):1995-2008. doi: 10.1056/NEJMoa1911303.

2: Solomon SD, McMurray JJV, Claggett B, de Boer RA, DeMets D, Hernandez AF, Inzucchi SE, Kosiborod MN, Lam CSP, Martinez F, Shah SJ, Desai AS, Jhund PS, Belohlavek J, Chiang CE, Borleffs CJW, Comin-Colet J, Dobreanu D, Drozd J, Fang JC, Alcocer-Gamba MA, Al Habeeb W, Han Y, Cabrera Honorio JW, Janssens SP, Katova T, Kitakaze M, Merkely B, O'Meara E, Saraiva JFK, Tereshchenko SN, Thierer J, Vaduganathan M, Vardeny O, Verma S, Pham VN, Wilderäng U, Zaozerska N, Bachus E, Lindholm D, Petersson M, Langkilde AM; DELIVER Trial Committees and Investigators. Dapagliflozin in Heart Failure with Mildly Reduced or Preserved Ejection Fraction. *N Engl J Med*. 2022 Sep 22;387(12):1089-1098. doi: 10.1056/NEJMoa2206286. Epub 2022 Aug 27. PMID: 36027570.

#### Data collection

Patients were randomized between February 15, 2017, and August 17, 2018, at 410 centers in 20 countries in the DAPA-HF trial, and between August 27, 2018, and December 30, 2020, at 353 centers in 20 countries in the DELIVER trial. Data were analyzed independently at the University of Glasgow.

#### Outcomes

Both trials were event-driven and had the same primary endpoint which was a composite of the time to the first occurrence of worsening heart failure or death from a cardiovascular cause. Worsening heart failure was defined as unplanned hospital admission for heart failure or an urgent visit for worsening heart failure resulting in the administration of an intravenous diuretic.

In this post-hoc paper, we analyzed a hierarchical composite outcome including the following components: all-cause death (Tier 1); ESKD or eGFR <15ml/min/1.73 m<sup>2</sup> (Tier 2); eGFR ≥57% decline (Tier 3); eGFR ≥50% decline (Tier 4); eGFR ≥40% decline (Tier 5); eGFR slope (Tier 6).

## Plants

#### Seed stocks

N/A

#### Novel plant genotypes

N/A

#### Authentication

N/A