nature portfolio

Peer Review File

Finerenone versus spironolactone in patients with chronic kidney disease and type 2 diabetes: a target trial emulation

Corresponding Author: Dr Vin-Cent Wu

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
This study provides real-world evidence comparing finerenone and spironolactone in patients with CKD and T2D. Its key strengths include the use of a robust methodology (propensity score matching on a large dataset, target trial emulation framework) to minimize confounding, a substantial matched cohort size (n=2,268), assessment of both efficacy (MACE, MAKE, mortality) and safety (hyperkalemia) outcomes, and demonstration of consistent benefits across multiple sensitivity and subgroup analyses. This offers evidence supporting finerenone's potential as a safer and more effective MRA option in this high-risk population within the studied timeframe. The data seem robust, however, there are some issues needed to be addressed.
Major:
1. The authors should avoid combining short-term medication users with long-term medication users into one group, which could cause dose-response bias. The treatment strategy should be predefined (e.g., "continuous medication use for ≥80% of the exposure period" versus "non-use"), rather than simply comparing "medication users" versus "non-users". In addition, please explicitly define "treatment discontinuation" (e.g., >30-day gap).
2. Table 2 reports only adjusted hazard ratios (aHR), limiting clinical interpretability. Absolute risk reduction and number needed to treat (NNT) for primary outcomes should be reported.
3. The results lack dose stratification (e.g., spironolactone 25mg vs. 50mg, finerenone 10mg vs. 20mg). Include dose subgroup analyses in the primary results.
4. The mechanistic explanation is insufficient: the study lacks in-depth exploration of how finerenone reduces mortality (e.g., "Is it mediated by reduced hyperkalemia → lower treatment discontinuation → improved long-term outcomes?").

Minor:
1. The authors should discuss how to weigh the efficacy gains of finerenone against its economic burden, particularly in low-resource settings, given that despite demonstrating survival benefits, its cost exceeds that of spironolactone by over 10 times.
2. Supplementary Table 3: Specify CPT codes for "acute dialysis".

Reviewer #2

(Remarks to the Author)
Thanks for the opportunity to review this manuscript, which reports a target trial emulation comparing finerenone to spironolactone in people with type 2 diabetes and chronic kidney disease. To my knowledge, this is the first study that attempts to directly compare between the two treatments among people with CKD and diabetes. The headline findings are that finerenone offered superior protection from death, cardiovascular or renal events compared to spironolactone, whilst also resulting in fewer adverse treatment events, such as hyperkalaemia.

The use of MRA in CKD is highly topical. We have recently seen the results of the FIDELITY pooled analysis of MRA in people with CKD and diabetes, demonstrating the potential cardiovascular and renal benefit of finerenone in people with CKD. In contrast, BARACK-D did not find that spironolactone improved cardiovascular outcomes in older people with stage 3b CKD, most of who did not have proteinuria. This study supports the outcomes of these two studies by suggesting finerenone offers improved efficacy compared to spironolactone.

The MACE event rate seems very high – 9.9% of those with finerenone and 13.3% among the spiro group over just 15m follow-up; that's more than double that reported in FIGARO. Could you comment on this?
Some geographical description of the continents/ countries covered within TriNetX would be helpful in understanding the context of the data. Is including both ethnicity and race distinctions instructive?

I think it would be helpful to include the fact that around 20% of participants had significant proteinuria based on the urine PCR results; this compares to around 2/3 of participants in FIDELITY. That could be emphasised more in the discussion of the cohort, in the comparison to the existing literature and in the discussion and conclusions.

Please can you report the baseline eGFR into more discrete categories ie separating CKD stage 3 into a and b? This should also ideally be consistent with figure 3; that currently compares between eGFR at 45 threshold whilst the comparison in table 1 is at a threshold of 30.

There is no discussion of missing data, but presumably this was an issue for at least some of the characteristics of interest such as BMI or lab results. Could you add information about how these data were handled in the analysis?

Line 293-295 (Methods, Target Trial Emulation) implies a full protocol is include in the supplementary material but I cannot see that. Does this refer to the full list of inclusion /exclusion criteria and if so could that be made clear? Presumably a protocol exists because it is stated that outcomes, covariates etc were pre-defined. Could you include where the protocol was published or upload it as a supplementary file?

Thank you for acknowledging the potential limitation of using routinely coded data for outcome ascertainment. I wonder if you could expand on this a little in terms of the choice of codes used? For example, the code list for MACE looks very focused – would this capture codes for events such as non-ST elevated MI or other MI sub-codes? Is there any precedent for the codes you used being used in prior research or how were they chosen?

You mention discontinuation in the limitations but are you able to report how many patients continued on treatment over follow-up and how this compared between the two groups? Can you add a limitation that the treatment dose appears to be unknown for a significant number of participants?

I'm not clear what you mean in this sentence – how does a higher potassium with finerenone reflect 'a more stable potassium balance' or 'compromise effectiveness'? Please can you re-write for clarity?
"Notably, even at baseline before matching, the finerenone group had higher potassium levels, suggesting a more stable potassium balance compared with spironolactone, potentially compromising effectiveness."

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
All of my previous comments have been addressed. I have no further comments.

Reviewer #2

(Remarks to the Author)
Thank you for submitting your revision and addressing the previous comments. I would recommend this is now accepted for publication. Congratulations on a well conducted study.

Open Access This Peer Review File 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 license, and indicate if changes were made.
In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.
The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit https://creativecommons.org/licenses/by/4.0/

Point-by-point response to Reviewers' Comments

Reviewer #1 (Remarks to the Author):

This study provides real-world evidence comparing finerenone and spironolactone in patients with CKD and T2D. Its key strengths include the use of a robust methodology (propensity score matching on a large dataset, target trial emulation framework) to minimize confounding, a substantial matched cohort size (n=2,268), assessment of both efficacy (MACE, MAKE, mortality) and safety (hyperkalemia) outcomes, and demonstration of consistent benefits across multiple sensitivity and subgroup analyses. This offers evidence supporting finerenone's potential as a safer and more effective MRA option in this high-risk population within the studied timeframe. The data seem robust, however, there are some issues needed to be addressed.

Response:

We sincerely thank the reviewer for the thoughtful and constructive summary. In the following responses, we have addressed the reviewer's concerns point by point and have revised the manuscript accordingly to strengthen its clarity and interpretability.

Major:

1. **The authors should avoid combining short-term medication users with long-term medication users into one group, which could cause dose-response bias. The treatment strategy should be predefined (e.g., "continuous medication use for ≥80% of the exposure period" versus "non-use"), rather than simply comparing "medication users" versus "non-users". In addition, please explicitly define "treatment discontinuation" (e.g., >30-day gap).**

Response:

In our primary analysis, we employed an **intention-to-treat (ITT)** approach to emulate the assignment process in a randomized clinical trial and to avoid immortal time bias. While we acknowledge the reviewer's suggestion to define treatment strategies based on sustained exposure (e.g., "continuous use for ≥80% of the exposure period"), such an approach could not be implemented due to platform

limitations. Specifically, TriNetX does not support dynamic exposure definitions (e.g., treatment discontinuation based on a >30-day gap) or per-protocol analyses that censor patients upon treatment discontinuation. To address this limitation within the constraints of the data, we examined treatment continuation rates as a proxy for persistence. To address this limitation, we examined treatment continuation rates as a proxy for real-world persistence: over the median 1.3-year follow-up, 49.0% of finerenone users and 55.5% of spironolactone users had ongoing prescriptions at 6 months, decreasing to 33.8% and 28.3% at 12 months, respectively. It is important to note that the numerator did not exclude patients who were censored (e.g., died or disenrolled), so true continuation proportions cannot be lower than reported.) Thus, persistence was modest in both arms and, if anything, slightly higher among finerenone users after the first year, making differential exposure an unlikely explanation for the observed risk estimates. These findings indicate generally modest persistence in both groups, with relatively higher continuation observed among finerenone users at later time points.

To explore the impact of sustained medication use, we conducted two complementary sensitivity analyses. To assess treatment persistence, we examined the proportion of patients who continued to receive prescriptions at 6 and 12 months after treatment initiation. As a sensitivity analysis, we performed landmark analyses at these timepoints, restricting the analysis to patients who remained event-free and had evidence of ongoing medication use up to the respective landmark. In the 3-month landmark cohort, finerenone was associated with lower risks of MAKE (aHR 0.51; 95% CI, 0.26–0.98) and mortality (aHR 0.26; 95% CI, 0.11–0.64), with a trend toward lower MACE (aHR 0.66; 95% CI, 0.41–1.08). In the 6-month cohort, risks of MACE (aHR 0.43; 95% CI, 0.19–0.97) and MAKE (aHR 0.33; 95% CI, 0.12–0.91) remained significantly lower, with a nonsignificant trend toward reduced mortality (aHR 0.16; 95% CI, 0.02–1.31). These results support the role of sustained treatment in maximizing benefit.

Taken together, these strategies help address concerns regarding treatment adherence and exposure misclassification to the extent possible within the limitations of the

available data. These results have been added to the **Results**, with corresponding methods and limitations detailed in the **Methods** and **Discussion** sections, respectively.

Supplement Figure 7

Outcomes	finerenone	spironolactone	aHR (95%CI)	P value
Landmark at 3 months (～3 months of sustained treatment) (n=368)				
MACE	27 (7.3%)	40 (10.9%)	0.66 (0.41-1.08)	0.095
MAKE	13 (3.5%)	25 (6.8%)	0.51 (0.26-0.98)	0.042
All-cause mortality	10 (2.7%)	23 (6.3%)	0.26 (0.11-0.64)	0.002
Landmark at 6 months (3–6 months of sustained treatment) (n=243)				
MACE	10 (4.1%)	20 (8.2%)	0.43 (0.19-0.97)	0.035
MAKE	10 (4.1%)	16 (6.6%)	0.33 (0.12-0.91)	0.024
All-cause mortality	10 (4.1%)	10 (4.1%)	0.16 (0.02-1.31)	0.051

Landmark analyses were conducted at 3- and 6-months post-index. The 3-month cohort comprised patients with continuous treatment and no outcome events during the first 90 days, with follow-up commencing on day 91. The 6-month cohort comprised patients with continuous treatment and no events between days 91 and 180, with follow-up commencing on day 181. Patient counts between 1 and 10 are reported as 10 to protect patient privacy.

Abbreviations: aHR, adjust hazard ratio; CI, confidence interval; MACE, major adverse cardiac events; MAKE, major adverse kidney events

Methods

To assess treatment persistence, we calculated the proportion of patients with ongoing prescriptions at 6 and 12 months after initiation. As a sensitivity analysis, we performed landmark analyses at these timepoints, including only patients who were event-free and remained on treatment to examine the associations with subsequent outcomes.

Results

Treatment persistence was modest overall, with 49.0% of patients remaining on finerenone and 55.5% on spironolactone at 6 months, declining to 33.8% and 28.3% at 12 months, respectively. In landmark analyses restricted to patients with continued treatment at 3 and 6 months, finerenone was consistently associated with lower risks of MACE, MAKE, and all-cause mortality compared with spironolactone (Supplement Table 7).

Discussion

Fourth, treatment adherence over time could not be fully assessed; exposure was defined at baseline, and dynamic changes such as discontinuation or switching were not captured. However, treatment persistence and landmark sensitivity analyses supported the robustness of our findings among sustained users, although residual misclassification was possible.

2. **Table 2 reports only adjusted hazard ratios (aHR), limiting clinical interpretability. Absolute risk reduction and number needed to treat (NNT) for primary outcomes should be reported.**

Response:

We thank the reviewer for highlighting the importance of absolute measures. In response, we have included absolute risk differences and numbers needed to treat (NNTs) with 95% confidence intervals for primary outcomes in **Table 2.** These additions enhance clinical relevance alongside adjusted hazard ratios. Relevant descriptions have also been added to the **Methods**, **Results**, and **Discussion** sections.

Table 2

Outcomes	No. of Patients with outcome		aHR (95%CI)	P-value	*E*-value (upper CI)	ARR (95% CI)	NNT (95% CI)
	Finerenone (n=1,132)	Spironolactone (n=1,132)					
MACE	112	150	0.74 (0.58-0.94)	0.013	2.06 (1.33)	0.03 (0.01-0.06)	29 (17-143)
MAKE	46	96	0.47 (0.33-0.67)	<0.001	3.68 (2.35)	0.04 (0.02-0.06)	23 (16-42)
All-cause mortality	35	112	0.31 (0.21-0.45)	<0.001	5.91 (3.83)	0.07 (0.05-0.09)	15 (11-21)

Methods

Absolute risk reduction (ARR) was calculated as the difference in event rates between groups, and number needed to treat (NNT) was derived as its reciprocal (NNT = 1/ARR). Confidence intervals for ARR and NNT were estimated using standard binomial methods.

Results

Comparison of Finerenone and Spironolactone on Outcome of Interests

After a follow-up of 1.3 years (IQR, 0.8-1.5), treatment with finerenone was associated with a significantly lower risk of all-cause mortality compared with spironolactone (aHR, 0.31; 95% CI, 0.21-0.45; P < 0.001). A total of 35 patients (3.1%) in the finerenone group and 112 patients (9.9%) in the spironolactone group died.

This corresponded to an absolute risk reduction (ARR) of 7% (95% CI, 5–9) and a number needed to treat (NNT) of 15 (95% CI, 11–21) (Table 2).

Finerenone was associated with a significantly lower risk of MACE than spironolactone (aHR, 0.74; 95% CI, 0.58–0.94; P = 0.013), with an ARR of 3% (95% CI, 1%–6%) and a NNT of 29 (95% CI, 17–143) (Table 2).

MAKE occurred in 46 patients (4.1%) in the finerenone group and 96 patients (8.5%) in the spironolactone group, with a lower hazard ratio in the finerenone group (aHR, 0.47; 95% CI, 0.33–0.67; P < 0.001). This corresponded to an ARR of 4% (95% CI, 2%–6%) and an NNT of 23 (95% CI, 16–42) (Table 2).

Discussion

The low NNT for mortality (15) from our study underscores the potential benefit even over short-term follow-up in this high-risk population. Furthermore, finerenone appears to act synergistically with RAS inhibitors, providing a more comprehensive strategy for managing cardiorenal complications in T2D.

3. **The results lack dose stratification (e.g., spironolactone 25mg vs. 50mg, finerenone 10mg vs. 20mg). Include dose subgroup analyses in the primary results.**

Response:

We appreciate the reviewer's suggestion to include dose-stratified analyses. Among those with available dose data, the vast majority were prescribed finerenone 10 mg or spironolactone 25 mg, while only a small proportion received finerenone 20 mg (5.6%) or spironolactone 50 mg (3.6%) prior to propensity score matching. These low frequencies limited the feasibility and statistical power of conducting meaningful stratified analyses across dose levels.

In response to the reviewer's comment, we conducted a dose-restricted sensitivity analysis comparing patients prescribed finerenone 10 mg with those receiving spironolactone 25 mg, representing the most common dosing patterns. The results were consistent with our primary findings, further supporting the robustness of our conclusions. In this dose-restricted cohort, finerenone was associated with significantly lower risks of MACE (aHR 0.70, 95% CI 0.52–0.93; p=0.013), MAKE (aHR 0.48, 95% CI 0.31–0.75; p<0.001), and all-cause mortality (aHR 0.35, 95% CI 0.22–0.55; p<0.001), consistent with primary analysis. Details of this dose-restricted sensitivity analysis are described in the **Methods** section and summarized in **Supplement Table 8**.

This observed prescribing pattern is consistent with our study population of patients with CKD and eGFR < 60 mL/min/1.73 m^2, and may reflect real-world clinical practice, in which lower starting doses are commonly used to mitigate the risk of hyperkalemia, despite the potential for greater antiproteinuric effects at higher doses. This is aligned with dosing strategies observed in pivotal randomized trials of MRAs, which also frequently utilized lower doses in this population.

We have now added this point to the **Discussion** section as a study limitation, noting that a significant proportion of patients had unknown dose data, which precluded more granular dose-response analyses.

Supplement Table 8

Included only patients with documented drug doses (10 mg for finerenone and 25 mg for spironolactone) (n=666)				
MACE	80 (12.0%)	109 (16.4%)	0.70 (0.52-0.93)	0.013
MAKE	30 (4.5%)	50 (7.5%)	0.48 (0.31-0.75)	<0.001
All-cause mortality	25 (3.8%)	70 (10.5%)	0.35 (0.22-0.55)	<0.001

Discussion

Fifth, the predominance of spironolactone 25 mg and finerenone 10 mg prescriptions limited our ability to evaluate dose–response effects. A dose-restricted sensitivity analysis confined to these common doses corroborated the primary findings. Moreover, a notable portion of patients had missing or unrecorded MRA dosing data, further restricting detailed dose stratification.

4. **The mechanistic explanation is insufficient: the study lacks in-depth exploration of how finerenone reduces mortality (e.g., "Is it mediated by reduced hyperkalemia → lower treatment discontinuation → improved long-term outcomes?").**

Response:

We thank the reviewer for the thoughtful comment regarding potential mechanisms underlying the observed mortality reduction with finerenone. We agree that reduced hyperkalemia risk—leading to fewer treatment discontinuations and more sustained exposure—may plausibly contribute to improved outcomes.

Our findings suggest that although early treatment persistence (~6 months) was slightly higher with spironolactone (55.5% vs. 49.0%), finerenone users demonstrated better continuation between 6 and 12 months (33.8% vs. 28.3%), indicating improved long-term tolerability. It is important to note that the numerator did not exclude patients who were censored (e.g., died or disenrolled), and therefore the actual discontinuation rates of spironolactone group should be higher than these estimates suggest. However, spironolactone use is often limited by early discontinuation, with nearly two-thirds stopping treatment within 6 months in the BARACK-D trial, compared to lower discontinuation rates (~27–29%) over full follow-up in finerenone trials. In our real-world study, treatment persistence between 6 and 12 months was higher in the finerenone group (33.8%) compared to spironolactone (28.3%). This pattern may reflect better long-term tolerability with finerenone.

This difference may be partially explained by the consistently lower rates of hyperkalemia observed with finerenone:

Serum potassium (finerenone vs spironolactone)

≥5.5 mmol/L: 17.2% vs. 26.4%; p-value< 0.001

≥6.0 mmol/L: 5.9% vs. 10.2%; p-value< 0.001

≥6.5 mmol/L: 2.1% vs. 3.9%; p-value= 0.014

These findings likely reflect finerenone's more selective pharmacologic profile, resulting in reduced potassium retention. While the dataset did not allow us to link hyperkalemia events directly to treatment discontinuation—a limitation we acknowledge—the lower hyperkalemia rates with finerenone may have facilitated more consistent exposure and contributed to improved survival. In large phase-III trials, finerenone produced durable cardiorenal and survival gains; its comparatively low hyperkalemia-driven discontinuation rate appears central to these outcomes.(Bakris, Agarwal et al. 2020) We also observed a higher rate of cardiac arrest in the spironolactone group, a known consequence of severe hyperkalemia, further supporting this hypothesis.

Furthermore, finerenone's additional anti-inflammatory and antifibrotic properties—described in prior studies(Alicic, Neumiller et al. 2025)—may also play a role. Due to limited availability of proteinuria and biomarker data, we were unable to explore these pathways directly and have now noted this limitation in the **Discussion**. We have revised the manuscript accordingly to reflect these mechanistic considerations and the need for further study.

Discussion

Both steroidal and non-steroidal MRAs mitigate endothelial dysfunction, oxidative stress, and albuminuria via aldosterone blockade.[8,17-19] Spironolactone has been shown to reduce left ventricular (LV) mass, improve arterial stiffness, and enhance LV function in early CKD.[20,21] However, spironolactone use is often limited by early discontinuation, with nearly two-thirds stopping treatment within 6 months in the BARACK-D trial,[10] compared to lower discontinuation rates (~27–29%) over full follow-up in finerenone trials.[4,5] In our real-world study, treatment persistence between 6 and 12 months was higher in the finerenone group (33.8%) compared to spironolactone (28.3%). This pattern may reflect better long-term tolerability in patients with finerenone.

Hyperkalemia remains a key limitation of MRAs therapy in CKD management.

Discontinuation or down-titration of MRAs, due to hyperkalemia is associated with significantly increased risks of adverse outcomes in patients with CKD and heart failure. The BIOSTAT-CHF and Swedish HF registries previously found that the hyperkalemia itself was not directly linked to adverse outcomes—instead, it marked suboptimal therapy that led to worse prognosis.[22,23] In our study, finerenone was associated with a lower risk of hyperkalemia compared with spironolactone, which aligns with prior research.[15] Before matching, mean serum potassium was marginally higher in participants treated with finerenone than in those given spironolactone, a disparity that likely reflects clinicians' preference for finerenone in patients with hyperkalemia liability. This imbalance was eliminated after matching, and finerenone still conferred a lower incidence of treatment-emergent hyperkalemia than spironolactone. Our specificity analyses yielded consistent findings, with finerenone being associated with greater benefit than spironolactone, particularly in the lower risk of cardiac arrest. This observation aligns with the lower incidence of hyperkalemia-related sudden cardiac death observed in the finerenone group. These findings support the potential role of finerenone as an alternative to spironolactone in the management of patients with CKD and type 2 diabetes, particularly among those at increased risk of potassium dysregulation.

Furthermore, the observed association between finerenone and reduced mortality and cardiorenal outcomes may be partly attributed to its distinct pharmacological properties beyond treatment persistence and lower hyperkalemia risk.[14] Finerenone has distinct pharmacological properties, including inhibition of pro-inflammatory and pro-fibrotic pathways via suppression of serum/glucocorticoid-regulated kinase 1 (Sgk1) activation and prevention of cofactor recruitment to mineralocorticoid receptors.[24-26] Finerenone shows superior antifibrotic effects in the heart and kidneys, slowing CKD progression in patients with diabetes.[15] These pharmacological distinctions provide a potential explanation for our findings and highlight finerenone's role as a more targeted therapeutic option in CKD and T2D.[24]

Additionally, the reasons for treatment discontinuation such as hyperkalemia or other adverse events were not captured in the database, limiting further insight into discontinuation patterns.

Minor:

1. **The authors should discuss how to weigh the efficacy gains of finerenone against its economic burden, particularly in low-resource settings, given that despite demonstrating survival benefits, its cost exceeds that of spironolactone by over 10 times.**

Response:

We thank the reviewer for raising this important and highly relevant clinical and policy question. We agree that weighing the survival and cardiorenal benefits of finerenone against its substantially higher cost—particularly in low-resource settings—is critical for real-world applicability and decision-making.

Due to the nature of the TriNetX platform, which contains only de-identified patient-level data, we were unable to access individual-level healthcare utilization or cost information, nor link across databases to perform a formal cost-effectiveness analysis. We acknowledge this as a limitation of our study and have now noted it in the **Discussion**.

Nonetheless, we would like to highlight that several modeling studies have evaluated the economic value of finerenone using established health economic frameworks. Notably, the FINE-CKD model(Cherney, Drzewiecka et al. 2025), along with other cost-effectiveness analyses conducted in the United States, the Netherlands, China, and Japan, have suggested that finerenone may be cost-effective in patients with CKD and T2D, particularly when long-term cardiovascular and kidney benefits are considered.(Quist, van Schoonhoven et al. 2023, Ming, Hong et al. 2024, Igarashi, Ohara et al. 2025, Zheng, Wu et al. 2025)

In contrast, the cost-effectiveness evidence for spironolactone in CKD remains limited. To our knowledge, the BARACK-D trial is the only major study to report economic outcomes, which estimated that the addition of low-dose spironolactone resulted in a cost per quality-adjusted life-year (QALY) gained exceeding the £30,000

threshold set by the National Institute for Health and Care Excellence (NICE), suggesting suboptimal cost-effectiveness under UK criteria.(Hobbs, McManus et al. 2024) Actually, evidence from U.S. and EU cost-utility models report incremental cost-effectiveness ratios (ICERs) of US $45 000–65 000 per QALY when finerenone is added to optimized care in CKD patients—figures that sit below or near commonly used willingness-to-pay thresholds.(Cherney, Drzewiecka et al. 2025, Zheng, Wu et al. 2025) The wholesale acquisition cost of finerenone (~ US $20 day) currently exceeds that of spironolactone (> 10-fold). In health systems where the WHO-recommended threshold is ≤ 1× GDP per capita per QALY, (https://repository.chds.hsph.harvard.edu/repository/2622/) such pricing would render finerenone non-viable.

Therefore, while finerenone is significantly more expensive than spironolactone (by more than tenfold), its broader clinical benefits may potentially offset the higher cost, especially in high-risk populations. That said, direct comparative cost-effectiveness analyses between finerenone and spironolactone remain scarce, and further research—including real-world studies incorporating cost data—is warranted to fully inform value-based decision-making in diverse healthcare settings. This point has been incorporated into the Discussion section.

Discussion

As finerenone moves toward broader clinical adoption, its higher cost relative to spironolactone warrants careful consideration. Although modeling studies suggest that finerenone may be cost-effective in CKD and T2D,[28-32] cost-effectiveness data for spironolactone remain limited.[10] Rigorous cost-utility analyses in resource-limited settings are therefore required before widespread adoption.

2. Supplementary Table 3: Specify CPT codes for "acute dialysis".

Response:

To avoid ambiguity, we have replaced the term **"acute dialysis"** with **"ever dialysis,"** denoting any inpatient dialysis event, whether transient or the prelude to chronic therapy. **Supplement Table 3** now lists the exact CPT codes used—90935 and 90947 (for single-session therapies)—along with other dialysis procedure codes for completeness. This inclusive definition acknowledges that recovery trajectories vary yet ensures every dialysis initiation is captured. The **Table** and **Methods** have been amended accordingly.

Supplement Table 3

#2: Major adverse kidney events (have any of the following)		
ESKD-related diagnoses		
Diagnosis	UMLS:ICD10CM:N18.6	End stage renal disease
Diagnosis	UMLS:ICD10CM:Z99.2	Dependence on renal dialysis
Dialysis initiation (includes acute dialysis procedures)		
Procedure	UMLS:CPT:90935	Hemodialysis procedure with single evaluation by a physician or other qualified health care professional (Acute dialysis)
Procedure	UMLS:CPT:90947	Dialysis procedure other than hemodialysis (eg, peritoneal dialysis, hemofiltration, or other continuous renal replacement therapies) requiring repeated evaluations by a physician or other qualified health care professional, with or without substantial revision of dialysis prescription (Acute dialysis)
Procedure	UMLS:CPT:90937	Hemodialysis procedure requiring repeated evaluation(s) with or without substantial revision of dialysis prescription
Procedure	UMLS:CPT:1006747	Hemodialysis Access, Intervascular Cannulation for Extracorporeal Circulation, or Shunt Insertion Procedures on Arteries and Veins
Procedure	UMLS:CPT:1012752	Hemodialysis Procedures
Procedure	UMLS:CPT:1012740	Dialysis Services and Procedures
Procedure	UMLS:SNOMED:302497006	Hemodialysis
Procedure	UMLS:HCPCS:C1750	Catheter, hemodialysis/peritoneal, long-term

Among dialysis procedure codes, CPT 90935 and 90947 were classified as acute dialysis, commonly used in the setting of acute kidney injury (AKI). Other codes were retained to capture dialysis initiation more broadly but are not specific to acute settings.

Methods

Exclusion criteria included prior eGFR values <15 mL/min/1.73m², end-stage renal disease (ESRD), or recent events such as acute coronary syndromes, stroke, cardiac arrest, cardiogenic shock, or ever dialysis within 60 days of the index prescription.

Reviewer #2 (Remarks to the Author):

Thanks for the opportunity to review this manuscript, which reports a target trial emulation comparing finerenone to spironolactone in people with type 2 diabetes and chronic kidney disease. To my knowledge, this is the first study that attempts to directly compare between the two treatments among people with CKD and diabetes. The headline findings are that finerenone offered superior protection from death, cardiovascular or renal events compared to spironolactone, whilst also resulting in fewer adverse treatment events, such as hyperkalaemia.

The use of MRA in CKD is highly topical. We have recently seen the results of the FIDELITY pooled analysis of MRA in people with CKD and diabetes, demonstrating the potential cardiovascular and renal benefit of finerenone in people with CKD. In contrast, BARACK-D did not find that spironolactone improved cardiovascular outcomes in older people with stage 3b CKD, most of who did not have proteinuria. This study supports the outcomes of these two studies by suggesting finerenone offers improved efficacy compared to spironolactone.

Response:

We thank the reviewer for the insightful summary and for recognizing the novelty and relevance of our study. We appreciate the contextualization of our findings within the broader landscape of MRA trials, including FIDELITY, and BARACK-D, and agree that our real-world evidence helps to inform the comparative effectiveness of finerenone versus spironolactone in patients with CKD and type 2 diabetes.

1. **The MACE event rate seems very high – 9.9% of those with finerenone and 13.3% among the spiro group over just 15m follow-up; that's more than double that reported in FIGARO. Could you comment on this?**

Response:

We thank the reviewer for highlighting this apparent discrepancy. In quick response, we agree that the MACE rate appears higher than that reported in FIGARO-DKD; however, this is expected given key differences in study populations. Our cohort closely parallels FIDELIO-DKD across key clinical parameters—including mean eGFR, distribution of CKD stage, and median UACR—whereas FIGARO-DKD deliberately enrolled patients with substantially earlier-stage CKD and lower baseline risk. The basis for this alignment is outlined in detail below.

Feature	Our real-world cohort	FIDELIO-DKD	FIGARO-DKD
Mean eGFR (mL min^{-1} 1.73 m^{-2})	46.2	44.3	67.8
eGFR < 30 % of patients	37.0%	38%	17.4%
Median UACR (mg/g)	NA	852	308
Heart-failure history	22.8%	13–17% (HF preserved EF allowed)	HFrEF excluded
RASi use at baseline	66.9%	≈98%	>99%

Differences from FIGARO-DKD

In finerenone user, our study observed a 9.9% incidence of MACE over a median follow-up of 1.3 years. In comparison, the FIGARO-DKD trial reported a 12.4% incidence of MACE over a median follow-up of 3.4 years in a relatively healthier population. Several key differences likely contributed to the higher short-term MACE incidence observed in our real-world cohort:

1. Patients were older on average (mean age 68.9 vs. 64.1 years)

2. A higher proportion had advanced CKD (eGFR <30 mL/min/1.73m^2: 37% vs. 17.4%)

3. 22.8% had a history of heart failure—a population excluded from FIGARO-DKD

4. Use of RAS inhibitors was lower (66.9% vs. >99%)

These factors indicate a substantially higher baseline cardiovascular risk in our cohort, which plausibly contributed to the elevated event rates over a shorter follow-up period.

Efficacy Context from FIDELIO-DKD

The FIDELIO-DKD trial enrolled a sicker population than FIGARO-DKD (mean eGFR 44.4 mL/min; 87% with UACR ≥300 mg/g) and demonstrated a 13.0% MACE rate and 7.7% all-cause mortality over a median 2.6-year follow-up. In comparison, our real-world study observed a 9.9% incidence of MACE and 3.1% mortality over a median follow-up of 1.3 years. Additionally, our cohort had more patients with heart failure (22.8% vs. excluded HFrEF in FIDELIO), lower RASi use (66.9% vs. 98%), further influencing outcome patterns. These findings reinforce the consistency of finerenone's cardiovascular efficacy across clinical settings, and support the comparative advantage observed in our real-world analysis against spironolactone.

Real-World Treatment Adherence and Persistence

Real-world medication persistence is frequently suboptimal; population studies show that nearly half of patients who initiate a mineralocorticoid-receptor antagonist discontinue therapy within five years, most often because of hyperkalemia or worsening renal function.(Zahir, Bonde et al. 2022) By contrast, protocol-driven randomized trials maintain high adherence through intensive site monitoring and mandatory titration visits, sharply limiting differential exposure between study arms.(Wiegand and Anderson 1982) In our electronic-health-record analysis, such rigorous oversight is absent, so early drop-off—particularly with spironolactone, which is less well tolerated—both elevates baseline cardiovascular risk and dilutes observed treatment exposure.(Secora, Shin et al. 2020) Early discontinuation—particularly of spironolactone—likely amplifies baseline risk yet attenuates observed benefit, rendering our intention-to-treat estimates conservative but unbiased.

Taken together, these considerations confirm that the cardiovascular event rates we report are biologically and methodologically consistent with contemporary evidence and underscore the external validity of our comparative findings. The corresponding additions have been incorporated into the revised **Discussion** section.

Discussion

Despite differences in baseline characteristics and follow-up duration, the 9.9% MACE incidence and 3.1% mortality we observed over 1.3 years are broadly consistent with those reported in the FIDELIO-DKD trial, reinforcing the external validity of our real-world findings.

2. **Some geographical description of the continents/ countries covered within TriNetX would be helpful in understanding the context of the data. Is including both ethnicity and race distinctions instructive?**

Response:

We thank the reviewer for this helpful suggestion. In response, we have added a description of the geographic distribution of contributing healthcare organizations within the TriNetX global network to the Methods section. TriNetX includes data from 21 countries spanning the Americas, Europe/Middle East/Africa (EMEA), and Asia-Pacific (APAC) regions, including Australia, Belgium, Brazil, Bulgaria, Estonia, France, Georgia, Germany, Ghana, Israel, Italy, Japan, Lithuania, Malaysia, Poland, Singapore, Spain, Taiwan, United Arab Emirates, the United Kingdom, and the United States. This broad coverage supports the generalizability of our findings while acknowledging that the distribution of patients varies by region and data source volume.(Palchuk, London et al. 2023) We also appreciate the reviewer's point regarding the distinction between race and ethnicity. In the TriNetX platform, race and ethnicity are captured separately when available, consistent with U.S. Office of Management and Budget standards. In **Table 1**, ethnicity was categorized as "Hispanic or Latino" or "Not Hispanic or Latino," while race included "White," "Black or African American," "Asian," and "Native Hawaiian," and included a "Unknown" category to retain completeness for each variable. We have updated the **Methods** and **Discussion** sections accordingly to clarify these points.

Methods

These HCOs span 21 countries across the Americas, Europe/Middle East/Africa (EMEA), and Asia-Pacific (APAC) regions, including the United States, United Kingdom, Germany, France, Israel, Japan, Taiwan, and Australia.

Race and ethnicity are recorded separately in TriNetX, consistent with clinical documentation standards in the United States and other participating regions.

3. **I think it would be helpful to include the fact that around 20% of participants had significant proteinuria based on the urine PCR results; this compares to around 2/3 of participants in FIDELITY. That could be emphasised more in the discussion of the cohort, in the comparison to the existing literature and in the discussion and conclusions.**

Response:

We thank the reviewer for pointing out this important aspect of cohort characterization. We apologize for any confusion. In response, we have revised **Table 1** to fully present all UPCR categories, including <30 mg/g and those with no available measurement.

Our cohort included approximately 20% of patients with significant proteinuria (UPCR ≥300 mg/g), which is substantially lower than the approximately two-thirds of participants with elevated albuminuria in the FIDELITY program. This discrepancy may partly reflect the high rate of missing UPCR data (57–86%) in our dataset, which likely led to an underestimation of the true prevalence of proteinuria.

As a sensitivity analysis, we restricted the cohort to patients with recorded UPCR (n=452) showed results consistent with the primary analysis: aHRs were 0.83 (95% CI, 0.56–1.23) for MACE, 0.56 (95% CI, 0.35–0.88) for kidney events, and 0.29 (95% CI, 0.17–0.48) for mortality (**Supplement Table 9**). These findings support the robustness of our results among patients with available proteinuria data.

Missing laboratory data, especially proteinuria (UPCR), is a common limitation of EHR studies. We now emphasize this in the **Discussion** to better contextualize our findings relative to existing clinical trial data. Corresponding revisions have also been made in the **Methods**, **Results**, and **Table**.

Table 1

UPCR, mg/g						
≥300 mg/g	350 (26.0)	955 (5.6)	0.528	232 (20.5)	243 (21.5)	0.024
30-300 mg/g	130 (9.7)	596 (3.5)	0.250	97 (8.6)	107 (9.5)	0.031
<30 mg/g	96 (7.1)	779 (4.6)	0.207	93 (8.2)	81 (7.2)	0.036
No measurement	769 (57.2)	14,639 (86.3)	0.588	710 (62.7)	701 (61.9)	0.017

Supplement Table 9

Cohort including complete UPCR measurements (n=452)				
MACE	47 (10.4%)	54 (12.0%)	0.83 (0.56-1.234)	0.364
MAKE	29 (6.4%)	49 (10.8%)	0.555 (0.35-0.88)	0.011
All-cause mortality	18 (4.0%)	62 (13.7%)	0.29 (0.17-0.48)	<0.001

Methods

To address missing laboratory data (e.g., BMI, HbA1c, SBP, lipids, and UPCR), we included a distinct "no measurement" category for each variable in the propensity score model.

We also restricted analyses to patients with complete laboratory data to assess the impact of missingness.

Results

Additional analyses performed before PSM, censoring events within the first 30 days of treatment initiation, restricting the cohort to those with documented drug doses (10 mg for finerenone, 25 mg for spironolactone), excluding patients who switched to other MRAs, and limited to patients with complete laboratory data also supported the robustness of our findings (Supplement Table 8 and 9).

Discussion

Sixth, missing laboratory data, especially proteinuria (UPCR), is a common limitation of EHR studies. High UPCR missingness may underestimate proteinuria prevalence and affect generalizability.

4. **Please can you report the baseline eGFR into more discrete categories ie separating CKD stage 3 into a and b? This should also ideally be consistent with figure 3; that currently compares between eGFR at 45 threshold whilst the comparison in table 1 is at a threshold of 30.**

Response:

We thank the reviewer for this helpful suggestion to refine the reporting of baseline kidney function. In response, we have revised Table 1 to present estimated glomerular filtration rate (eGFR) using more discrete CKD categories, specifically by separating CKD stage 3a (≥45 mL/min/1.73m²) and 3b (30–44 mL/min/1.73m²), and retaining stage 4–5 (<30 mL/min/1.73m²) as a separate group. This categorization is now consistent across **Table 1** and **Figure 3**.

The updated data after propensity score matching show good covariate balance across eGFR strata, with all standardized mean differences (SMDs) <0.1, indicating adequate matching quality. The matched cohorts were well distributed across CKD stages, with 15~17% having eGFR ≥45, ~46% in the 30–44 range, and ~37–39% with eGFR <30.

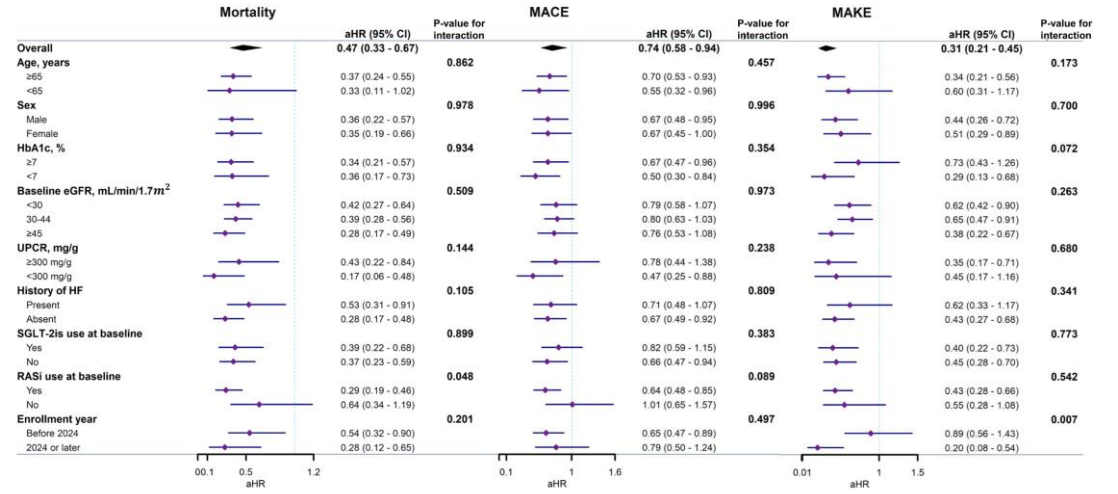
We also incorporated these updated eGFR categories in the subgroup analyses (Figure 3), which consistently showed no statistically significant heterogeneity in treatment effects across CKD stages (p for interaction >0.05 for all outcomes), indicating robust treatment effects irrespective of baseline renal function.

The corresponding sections in the **Results** have been updated to incorporate these changes.

Table 1

eGFR, mL/min/1.73m²						
≥45 mL/min/1.73m²	212 (15.8)	4,185 (24.7)	0.223	191 (16.9)	174 (15.4)	0.041
30-44 mL/min/1.73m²	616 (45.8)	7,348 (43.3)	0.050	521 (46.0)	521 (46.0)	<0.001
<30 mL/min/1.73m²	517 (38.4)	5,436 (32.0)	0.134	420 (37.1)	437 (38.6)	0.031

Figure 3



Results

The association between finerenone use and a lower risk of the primary outcomes remained consistent across subgroups stratified by age, sex, HbA1c level (≥7% vs <7%), kidney function (eGFR ≥45, 30–44, and <30 mL/min/1.73 m²), proteinuria (UPCR ≥300 vs <300 mg/g), heart failure status, concurrent SGLT2 inhibitor use, RAS inhibitor use, and enrollment year.

5. **There is no discussion of missing data, but presumably this was an issue for at least some of the characteristics of interest such as BMI or lab results. Could you add information about how these data were handled in the analysis?**

Response:

We acknowledge that several baseline covariates had missing values, including BMI, HbA1c, systolic blood pressure, lipid profile, and UPCR—a common limitation in real-world electronic health record research. The TriNetX platform addresses missingness in propensity-score modeling through a built-in function that assigns a dedicated "no measurement" category for each variable with missing data.(Palchuk, London et al. 2023) This approach retains the full cohort while allowing the propensity-score algorithm to account for differential ascertainment across treatment groups. The proportion of patients with missing data ("no measurement") for each variable has now been added to **Table 1**.

Because the TriNetX analytic engine does not support multiple imputation or alternative missing-data handling methods, we assessed the robustness of this missing-data strategy by repeating all primary analyses in a subset of patients with complete laboratory data. Although the smaller sample size led to wider confidence intervals, the point estimates remained directionally consistent with those of the main analysis (**Supplement Table 9**), supporting the validity of our approach.

The **Methods**, **Results**, and **Discussion** sections have been updated to describe our handling of missing data and to report the findings from these sensitivity analyses.

Supplement Table 9

Outcomes	finerenone	spironolactone	aHR (95%CI)	P value
Cohort including complete BMI, HbA1c, and SBP measurements (n=710)				
MACE	79 (11.1%)	113 (15.9%)	0.65 (0.49–0.86)	0.003
MAKE	38 (5.4%)	66 (9.3%)	0.55 (0.37–0.81)	0.003
All-cause mortality	31 (4.4%)	73 (10.3%)	0.40 (0.27–0.61)	<0.001
Cohort including complete BMI, A1c, SBP, and lipid profile measurements (n=552)				
MACE	61 (11.1%)	61 (11.1%)	0.81 (0.58–1.14)	0.222
MAKE	30 (5.4%)	49 (8.9%)	0.59 (0.37–0.93)	0.021
All-cause mortality	24 (4.3%)	52 (9.4%)	0.45 (0.28–0.72)	0.001
Cohort including complete UPCR measurements (n=452)				
MACE	47 (10.4%)	54 (12.0%)	0.83 (0.56-1.234)	0.364
MAKE	29 (6.4%)	49 (10.8%)	0.555 (0.35-0.88)	0.011
All-cause mortality	18 (4.0%)	62 (13.7%)	0.29 (0.17-0.48)	<0.001

Methods

To address missing laboratory data (e.g., BMI, HbA1c, SBP, lipids, and UPCR), we included a distinct “no measurement” category for each variable in the propensity score model.

We also restricted analyses to patients with complete laboratory data to assess the impact of missingness.

Results

Additional analyses performed before PSM, censoring events within the first 30 days of treatment initiation, restricting the cohort to those with documented drug doses (10 mg for finerenone, 25 mg for spironolactone), excluding patients who switched to other MRAs, and limited to patients with complete laboratory data also supported the robustness of our findings (Supplement Table 8 and 9).

Discussion

Sixth, missing laboratory data, especially proteinuria (UPCR), is a common limitation of EHR studies. High UPCR missingness may underestimate proteinuria prevalence and affect generalizability.

6. **Line 293-295 (Methods, Target Trial Emulation) implies a full protocol is include in the supplementary material but I cannot see that. Does this refer to the full list of inclusion /exclusion criteria and if so could that be made clear? Presumably a protocol exists because it is stated that outcomes, covariates etc were pre-defined. Could you include where the protocol was published or upload it as a supplementary file?**

Response:

We apologize for any confusion. The term "full protocol" in the **Methods** section refers to the detailed target trial specification as summarized in **Supplement Table 1** and **Supplement Figure 1**, which includes eligibility criteria, treatment strategies, outcome definitions, and analysis plans. We have revised the manuscript to clarify this point.

Methods

Details of the target trial specification, including eligibility criteria, treatment strategies, outcomes, and analysis approach, are provided in Supplement Table 1 and Supplement Figure 1.

Supplement Table 1

Note: This table serves as the complete protocol specification for the emulated target trial, following the framework of Hernán and Robins. All eligibility criteria, treatment strategies, outcome definitions, and analytic methods were pre-defined prior to data analysis. A separate protocol document was not created or published.

7. **Thank you for acknowledging the potential limitation of using routinely coded data for outcome ascertainment. I wonder if you could expand on this a little in terms of the choice of codes used? For example, the code list for MACE looks very focused – would this capture codes for events such as non-ST elevated MI or other MI sub-codes? Is there any precedent for the codes you used being used in prior research or how were they chosen?**

Response:

We thank the reviewer for the thoughtful comment regarding the specificity of our outcome definitions and the use of routinely coded data. Actually, we used a broader ICD-10-CM category such as I21 (acute myocardial infarction), all subordinate codes—including I21.4 for non–ST elevation myocardial infarction (NSTEMI) and I21.x for ST elevation myocardial infarction (STEMI). In TriNetX, diagnosis codes are organized within a hierarchical structure based on the Unified Medical Language System (UMLS). This approach allows comprehensive capture of clinically relevant subtypes within each category and enhances outcome ascertainment consistency.

In our original manuscript, MACE was defined as acute myocardial infarction, nonfatal stroke, hemorrhagic stroke, cardiac arrest, and cardiogenic shock. However, based on our actual code selection (primarily I21 and I24), a more accurate description is that we defined MACE to include acute coronary syndromes (ACS) rather than only AMI. To improve clarity and consistency, we have revised the **Methods** and **Supplementary Materials** to reflect this coding framework, and we have updated the terminology throughout the manuscript accordingly. Our outcome construction is aligned with standard definitions of MACE used in large cardiovascular studies. Moreover, a systematic review by Bosco et al. (2021) highlighted the heterogeneity in MACE definitions across EHR and claims-based studies, supporting the validity of clinically oriented, hierarchically structured code groupings such as ours.(Bosco, Hsueh et al. 2021)

To further address the reviewer's concern regarding diagnostic specificity, we conducted additional component analyses using narrower definitions of myocardial

infarction and MACE. Specifically, we estimated hazard ratios for: (1) AMI only (ICD-10-CM I21), (2) STEMI only (I21.0–I21.3), (3) MACE defined using AMI (I21) instead of ACS (I21 + I24), and (4) MACE defined using STEMI (I21.0–I21.3) instead of ACS. These definitions were applied consistently across all MACE components (cardiac arrest, cardiogenic shock, stroke).

Across all alternatives, hazard-ratio estimates remained directionally consistent with the primary analysis (**Supplementary Table 5**), confirming the robustness of our findings. We have updated the **Methods, Results, Discussion, and Supplementary Materials** to reflect these additional analyses and to acknowledge the inherent limitations of routinely coded administrative data, including potential misclassification and coding heterogeneity.

Supplement Table 3

Note: In TriNetX, diagnosis codes are mapped hierarchically. Selecting a parent code (e.g., I21) automatically includes all associated sub-codes (e.g., I21.4 for NSTEMI), ensuring complete inclusion of relevant diagnostic subtypes.

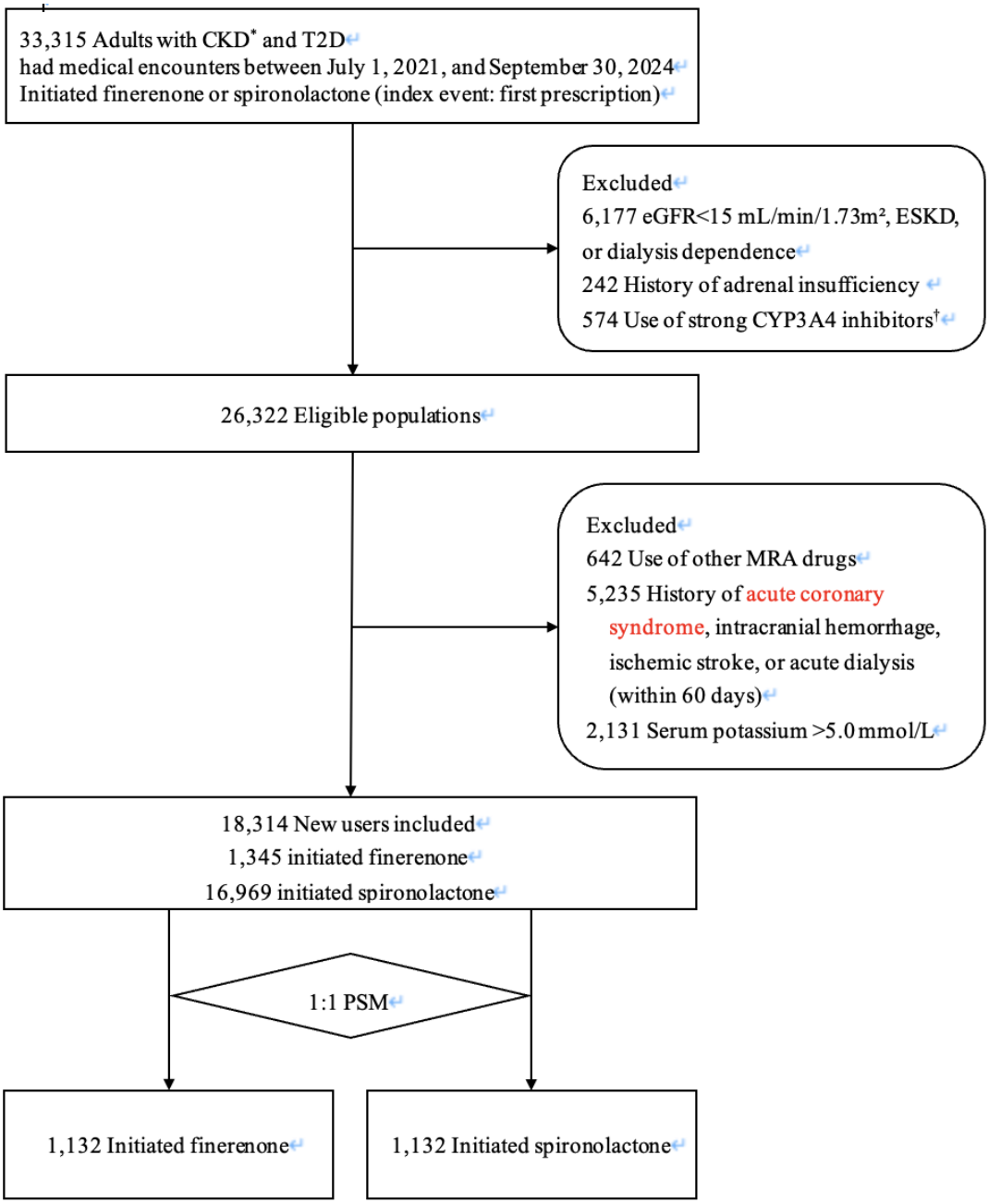
Supplement Table 5

Supplement Table 5. Adjusted hazard ratios for components and alternative definitions of major adverse cardiovascular events

Outcomes	finerenone (n=1,132)	spironolactone (n=1,132)	aHR (95%CI)	P value
MACE using ACS	112 (9.9%)	150 (13.3%)	0.74 (0.58-0.94)	0.013
ACS	64 (5.7%)	83 (7.3%)	0.77 (0.55-1.06)	0.108
AMI only	57 (5.0%)	78 (6.9%)	0.691 (0.51-0.94)	0.019
STEMI only	14 (1.2%)	18 (1.6%)	0.743 (0.40,1.4)	0.357
Cardiac arrest/ Cardiogenic shock	10 (0.9%)	24 (2.1%)	0.33 (0.15-0.74)	0.005
Ischemic Stroke	49 (4.3%)	52 (4.6%)	0.93 (0.63-1.38)	0.729
Nontraumatic hemorrhage	10 (0.9%)	19 (1.7%)	0.48 (0.22-1.05)	0.061
MACE using AMI only	108 (9.5%)	149 (13.2%)	0.68 (0.54-0.85)	<0.001
MACE using STEMI only	75 (6.6%)	102 (9.0%)	0.70 (0.53-0.91)	0.008

Abbreviations:
ACS, acute coronary syndromes (ICD-10-CM I21 + I24)
AMI, acute myocardial infarction (ICD-10-CM I21)
STEMI, ST-elevation myocardial infarction (ICD-10-CM I21.0–I21.3)
aHR, adjusted hazard ratio; CI, confidence interval; MACE, major adverse cardiovascular events

Figure 1


33,315 Adults with CKD* and T2D
had medical encounters between July 1, 2021, and September 30, 2024
Initiated finerenone or spironolactone (index event: first prescription)
Excluded
6,177 eGFR<15 mL/min/1.73m², ESKD, or dialysis dependence
242 History of adrenal insufficiency
574 Use of strong CYP3A4 inhibitors†
26,322 Eligible populations
Excluded
642 Use of other MRA drugs
5,235 History of acute coronary syndrome, intracranial hemorrhage, ischemic stroke, or acute dialysis (within 60 days)
2,131 Serum potassium >5.0 mmol/L
18,314 New users included
1,345 initiated finerenone
16,969 initiated spironolactone
1:1 PSM
1,132 Initiated finerenone
1,132 Initiated spironolactone

Methods

Exclusion criteria included prior eGFR values <15 mL/min/1.73m², end-stage renal disease (ESRD), or recent events such as acute coronary syndromes, stroke, cardiac arrest, cardiogenic shock, or ever dialysis within 60 days of the index prescription.

MACE was defined as acute coronary syndromes, nonfatal stroke, hemorrhagic stroke, cardiac arrest, cardiogenic shock.

Results

This association remained consistent across individual components, including acute myocardial infarction, cardiac arrest/ cardiogenic shock, and alternative MACE definitions (Supplement Table 5).

Discussion

Second, the reliance on EHRs and standardized coding systems, including the International Classification of Diseases, 10th Revision (ICD-10), may introduce ascertainment bias due to potential misclassification or incomplete capture of events. While negative control outcome was used to assess potential systemic bias, overdiagnosis, underdiagnosis, or misclassification cannot be fully excluded.

8. **You mention discontinuation in the limitations but are you able to report how many patients continued on treatment over follow-up and how this compared between the two groups? Can you add a limitation that the treatment dose appears to be unknown for a significant number of participants?**

Response:

We thank the reviewer for highlighting the importance of treatment continuation and dose information. To address this, we conducted a treatment persistence analysis to provide insight into longitudinal medication use. Specifically, we assessed the proportion of patients with evidence of continued prescriptions at fixed time intervals (e.g., ~6 and 6–12 months). Between months 6 and 12, treatment persistence was modest overall, but numerically higher in the finerenone group (33.8%) compared with spironolactone (28.3%), suggesting slightly better sustained use of finerenone in real-world practice. It is important to note that the numerator did not exclude patients who were censored (e.g., died or disenrolled), and therefore the actual continuation rates should be at least as high as these estimates suggest. These findings indicate generally modest persistence in both groups, with relatively higher continuation observed among finerenone users at later time points.

Among patients with available dose data, the vast majority received finerenone 10 mg or spironolactone 25 mg, while finerenone 20 mg (5.6%) and spironolactone 50 mg (3.6%) were infrequently used before propensity score matching. To address this, we performed a sensitivity analysis restricted to patients with documented prescriptions of finerenone 10 mg versus spironolactone 25 mg. The direction and magnitude of the treatment effect estimates remained consistent with the primary analysis, further supporting the robustness of our findings (**Supplement Table 8**).

Corresponding revisions have been made in the **Methods, Results**, and **Discussion** sections to address these limitations related to treatment persistence and missing dose information.

Supplement Table 8

Included only patients with documented drug doses (10 mg for finerenone and 25 mg for spironolactone) (n=666)				
MACE	80 (12.0%)	109 (16.4%)	0.70 (0.52-0.93)	0.013
MAKE	30 (4.5%)	50 (7.5%)	0.48 (0.31-0.75)	<0.001
All-cause mortality	25 (3.8%)	70 (10.5%)	0.35 (0.22-0.55)	<0.001

Methods

To assess treatment persistence, we calculated the proportion of patients with ongoing prescriptions at 6 and 12 months after initiation. As a sensitivity analysis, we performed landmark analyses at these timepoints, including only patients who were event-free and remained on treatment to examine the associations with subsequent outcomes.

Results

Treatment persistence was modest overall, with 49.0% of patients remaining on finerenone and 55.5% on spironolactone at 6 months, declining to 33.8% and 28.3% at 12 months, respectively. In landmark analyses restricted to patients with continued treatment at 3 and 6 months, finerenone was consistently associated with lower risks of MACE, MAKE, and all-cause mortality compared with spironolactone (Supplement Table 7).

Discussion

Fifth, the predominance of spironolactone 25 mg and finerenone 10 mg prescriptions limited our ability to evaluate dose–response effects. A dose-restricted sensitivity analysis confined to these common doses corroborated the primary findings. Moreover, a notable portion of patients had missing or unrecorded MRA dosing data, further restricting detailed dose stratification.

9. **I'm not clear what you mean in this sentence – how does a higher potassium with finerenone reflect 'a more stable potassium balance' or 'compromise effectiveness'? Please can you re-write for clarity?**
   "Notably, even at baseline before matching, the finerenone group had higher potassium levels, suggesting a more stable potassium balance compared with spironolactone, potentially compromising effectiveness."

Response:

We appreciate the reviewer's request for clarification and rewritten the sentences.

"Before matching, mean serum potassium was marginally higher in participants treated with finerenone than in those given spironolactone, a disparity that likely reflects clinicians' preference for finerenone in patients with hyperkalemia liability. This imbalance was eliminated after matching, and finerenone still conferred a lower incidence of treatment-emergent hyperkalemia than spironolactone."

The revised wording has been inserted into the **Discussion**, and the accompanying text now notes that the pre-matching difference disappeared after matching, underscoring adequate baseline balance in the analytic cohort.

Discussion

Before matching, mean serum potassium was marginally higher in participants treated with finerenone than in those given spironolactone, a disparity that likely reflects clinicians' preference for finerenone in patients with hyperkalemia liability. This imbalance was eliminated after matching, and finerenone still conferred a lower incidence of treatment-emergent hyperkalemia than spironolactone.

In addition, we made minor general corrections and editorial improvements to enhance clarity, consistency, and readability throughout the manuscript, even though these were not explicitly requested by the reviewers.

General Corrections / Editorial Changes

1. In the **Supplement Table 8**, the adjusted hazard ratio (aHR) and 95% confidence interval for the negative control outcome (cancer) was incorrectly reported as 1.82 (0.52–1.30) and has been corrected to 0.82 (0.52–1.33).

2. We corrected the title of **Supplement Table 5** to accurately reflect the statistical measures used. In the original version, the table was titled: **"Adjusted hazard ratios for components of major adverse cardiovascular events and for hyperkalemia at various serum potassium thresholds."** However, the estimates for hyperkalemia were derived from logistic regression models and thus reported as adjusted odds ratios rather than hazard ratios. To clarify this distinction, we have revised the presentation as follows: **"Adjusted hazard ratios for components and alternative definitions of major adverse cardiovascular events"** and **"Supplement Table 6. Adjusted odds ratios for hyperkalemia at various serum potassium thresholds"**.

3. In the **Discussion** section, the phrase "In a post hoc analysis of the FIDELITY study" has been revised to "In a post hoc analysis of the FIDELITY-TRH and AMBER studies" for accuracy.

Supplement Table 8

All-cause mortality	25 (2.6%)	105 (11.1%)	0.23 (0.15-0.36)	<0.001
Negative Outcome Control (n=800)				
Overall cancer	34 (4.3%)	41 (5.1%)	0.82 (0.52-1.33)	0.391

Supplement Table 5. Adjusted hazard ratios for components and alternative definitions of major adverse cardiovascular events

Outcomes	finerenone (n=1,132)	spironolactone (n=1,132)	aHR (95%CI)	P value
MACE using ACS	112 (9.9%)	150 (13.3%)	0.74 (0.58-0.94)	0.013
ACS	64 (5.7%)	83 (7.3%)	0.77 (0.55-1.06)	0.108
AMI only	57 (5.0%)	78 (6.9%)	0.691 (0.51-0.94)	0.019
STEMI only	14 (1.2%)	18 (1.6%)	0.743 (0.40,1.4)	0.357
Cardiac arrest/ Cardiogenic shock	10 (0.9%)	24 (2.1%)	0.33 (0.15-0.74)	0.005
Ischemic Stroke	49 (4.3%)	52 (4.6%)	0.93 (0.63-1.38)	0.729
Nontraumatic hemorrhage	10 (0.9%)	19 (1.7%)	0.48 (0.22-1.05)	0.061
MACE using AMI only	108 (9.5%)	149 (13.2%)	0.68 (0.54-0.85)	<0.001
MACE using STEMI only	75 (6.6%)	102 (9.0%)	0.70 (0.53-0.91)	0.008

Supplement Table 6. Adjusted odds ratios for hyperkalemia at various serum potassium thresholds

Outcomes	finerenone (n=1,132)	spironolactone (n=1,132)	OR (95%CI)	P value
Hyperkalemia				
≥5.5 mmol/L	195 (17.2%)	299 (26.4%)	0.58 (0.47-0.71)	<0.001
≥6.0 mmol/L	47 (5.9%)	115 (10.2%)	0.56 (0.41-0.76)	<0.001
≥6.5 mmol/L	24 (2.1%)	44 (3.9%)	0.54 (0.32-0.89)	0.014

Discussion

In a post hoc analysis of the FIDELITY-TRH and AMBER studies

Additional Reference

Alicic, R. Z., J. J. Neumiller and K. R. Tuttle (2025). "Combination therapy: an upcoming paradigm to improve kidney and cardiovascular outcomes in chronic kidney disease." Nephrol Dial Transplant **40**(Supplement_1): i3-i17.

Bakris, G. L., R. Agarwal, S. D. Anker, B. Pitt, L. M. Ruilope, P. Rossing, P. Kolkhof, C. Nowack, P. Schloemer, A. Joseph, G. Filippatos and F.-D. Investigators (2020). "Effect of Finerenone on Chronic Kidney Disease Outcomes in Type 2 Diabetes." N Engl J Med **383**(23): 2219-2229.

Bosco, E., L. Hsueh, K. W. McConeghy, S. Gravenstein and E. Saade (2021). "Major adverse cardiovascular event definitions used in observational analysis of administrative databases: a systematic review." BMC Med Res Methodol **21**(1): 241.

Cherney, D., A. Drzewiecka, K. Folkerts, P. Levy, A. Millier, S. Morris, M. Pochopien, P. Roy-Chaudhury, S. D. Sullivan and P. Mernagh (2025). "Cost-effectiveness of finerenone therapy for patients with chronic kidney disease and type 2 diabetes in England & Wales: results of the FINE-CKD model." J Med Econ **28**(1): 196-206.

Hobbs, F. D. R., R. J. McManus, C. J. Taylor, N. R. Jones, J. K. Rahman, J. Wolstenholme, S. Kim, J. Kwon, L. Jones, J. A. Hirst, L.-M. Yu, S. Mort, F. D. R. Hobbs, R. J. McManus, L. Jones, B. Thompson, J. K. Rahman, C. Vicary, L. Evans, E. Egden, M. Patil, L.-M. Yu, S. Mort, J. Wolstenholme, D. Lasserson, C. J. Taylor, N. R. Jones, J. Townend, C. Ferro, P. Bower, A. Farmer, D. Fitzmaurice, G. Feder, P. Little, N. Qureshi, F. D. R. Hobbs, R. Perera, D. Timmins, D. Fitzmaurice, G. Heer, R. Della, H. Duffy, F. McRonald, D. Popoola, K. Jheeta, G. Feder, S. Bryant, M. Taal, Y. Newey, D. Morgan, P. Bower, C. Gardner, V. Lee, T. Blakeman, N. Qureshi, L. Cross-Bardell, C. Brindley, P. Little, J. Barnett, K. Middleton, B.-D. I. on behalf of the, B.-D. Investigators, t. Regional coordinating centre, Oxford, Birmingham, Bristol, Derby, Manchester, Nottingham and Southampton (2024). "Low-dose spironolactone and cardiovascular outcomes in moderate stage chronic kidney disease: a randomized controlled trial." Nature Medicine **30**(12): 3634-3645.

Igarashi, A., K. Ohara, H. Matsuda, J. Morii, S. Jagannathan and R. Filomeno (2025). "Cost-Effectiveness Analysis of Finerenone for Treatment of Chronic Kidney Disease in Patients with Type 2 Diabetes from Japanese Payer Perspective." Adv Ther **42**(2): 995-1008.

Ming, J., G. Hong, Y. Xu, P. Mernagh, M. Pochopień and H. Li (2024). "Cost-effectiveness of Finerenone in Addition to Standard of Care for Patients with Chronic Kidney Disease and Type 2 Diabetes in China." Adv Ther **41**(8): 3138-3158.

Palchuk, M. B., J. W. London, D. Perez-Rey, Z. J. Drebert, J. P. Winer-Jones, C. N. Thompson, J. Esposito and B. Claerhout (2023). "A global federated real-world data

and analytics platform for research." JAMIA Open **6**(2): ooad035.
Quist, S. W., A. V. van Schoonhoven, S. J. L. Bakker, M. Pochopień, M. J. Postma, J. M. T. van Loon and J. H. J. Paulissen (2023). "Cost-effectiveness of finerenone in chronic kidney disease associated with type 2 diabetes in The Netherlands." Cardiovasc Diabetol **22**(1): 328.
Secora, A. M., J. I. Shin, Y. Qiao, G. C. Alexander, A. R. Chang, L. A. Inker, J. Coresh and M. E. Grams (2020). "Hyperkalemia and Acute Kidney Injury with Spironolactone Use Among Patients with Heart Failure." Mayo Clin Proc **95**(11): 2408-2419.
Wiegand, R. D. and R. E. Anderson (1982). "Determination of molecular species of rod outer segment phospholipids." Methods Enzymol **81**: 297-304.
Zahir, D., A. Bonde, C. Madelaire, M. Malmborg, J. H. Butt, E. Fosbol, G. Gislason, C. Torp-Pedersen, C. Andersson, P. Rossignol, J. J. V. McMurray, L. Kober and M. Schou (2022). "Temporal trends in initiation of mineralocorticoid receptor antagonists and risk of subsequent withdrawal in patients with heart failure: a nationwide study in Denmark from 2003-2017." Eur J Heart Fail **24**(3): 539-547.
Zheng, C., J. Wu, N. Li, X. Wei, Z. Huang, L. Chen and Z. Chen (2025). "Cost-effectiveness of finerenone added to standard of care for patients with type 2 diabetes-related chronic kidney disease in the United States." Diabetes Obes Metab **27**(1): 165-173.

Point-by-point response to Reviewers' Comments

Reviewer #1 (Remarks to the Author):

All of my previous comments have been addressed. I have no further comments.

Response:

We sincerely thank the reviewer for their careful review and constructive comments throughout the revision process. We appreciate the positive feedback and are glad that our revisions have addressed all concerns.

Reviewer #2 (Remarks to the Author):

Thank you for submitting your revision and addressing the previous comments. I would recommend this is now accepted for publication. Congratulations on a well conducted study.

Response:

We sincerely thank the reviewer for their thoughtful evaluation and kind recommendation for acceptance. We appreciate their positive comments on our study and are grateful for the opportunity to improve the manuscript.