

A meta-analysis of albuminuria as a surrogate endpoint for kidney failure

Received: 27 June 2025

Accepted: 14 October 2025

Published online: 06 November 2025

 Check for updates

Hiddo J. L. Heerspink  , Willem H. Collier^{3,4}, Juhi Chaudhari⁵, Shiyuan Miao⁵, Hocine Tighiouart^{6,7}, Gerald B. Appel⁸, Fernando Caravaca-Fontán , Jürgen Floege¹⁰, Thierry Hannedouche¹¹, Enyu Imai¹², Tazeen H. Jafar , Julia B. Lewis¹⁴, Philip K. T. Li , Francesco Locatelli¹⁶, Bart D. Maes , Brendon L. Neuen², Vlado Perkovic², Ronald D. Perrone⁵, Giuseppe Remuzzi , Francesco P. Schena¹⁹, Christoph Wanner , Tom Greene³ & Lesley A. Inker⁵

Albuminuria is a central biomarker in chronic kidney disease (CKD), used for the detection and prognosis of the disease. In clinical trials assessing CKD progression, change in the level of albuminuria is a candidate surrogate endpoint for kidney failure. Evaluation of the validity of this surrogate endpoint across a diverse range of interventions and populations is required to support its further acceptance. Here, in an individual participant data analysis of 48 randomized controlled trials (studies) involving 85,681 participants, we assessed the association between treatment effects on 6-month urinary albumin:creatinine ratio (UACR) change and the established clinical endpoint of kidney failure or doubling of serum creatinine concentrations. Across all trials, each 30% reduction in the geometric mean of the UACR in the treatment group relative to the control group was associated with an average of 19% lower hazard for the clinical endpoint (95% Bayesian credible interval (BCI): 5–30%); median coefficient of determination (R^2) = 0.66 (95% BCI: 0.06–0.98). There was no clear evidence that this association varied by CKD etiology. These results provide further support for use of albuminuria change as a surrogate endpoint in CKD clinical trials.

CKD is a global health problem affecting approximately 840 million people¹. New therapies to slow the progression of CKD have become available and are recommended by clinical practice guidelines². Despite these new therapies, a residual risk of kidney failure and cardiovascular complications persists in a substantial number of patients. Novel treatments or combinations of existing drugs are needed to halt the progression of CKD.

The availability of new therapeutic approaches to slow the progression of CKD has increased the emphasis on early recognition of CKD because appropriate treatments in the initial stages of disease are likely to be more beneficial than interventions applied at later stages³. However, efficacy assessment of therapies in early stages of CKD is challenging because the established clinical outcomes—kidney

failure or halving of glomerular filtration rate (GFR)—are typically observed within feasible clinical trial timeframes only for patients with advanced disease. Surrogate outcomes can help to assess drug efficacy when clinical outcome trials are not feasible⁴. Change in albuminuria is a biologically plausible surrogate endpoint for the progression of CKD. Experimental evidence suggests that increased albumin leakage through the glomerular basement membrane promotes inflammation and causes tubulointerstitial damage^{5–7}. There is strong and robust epidemiological evidence that the albuminuria level and changes in albuminuria over time are associated with kidney failure⁸.

The highest evidence to support validation of surrogate endpoints comes from trial-level meta-analyses that characterize association between treatment effects on the surrogate and clinical endpoints

Table 1 | Baseline disease characteristics overall and stratified by disease

| Group | Number of studies (participants) | Age (years) | Female | Diabetes | GFR (ml min ⁻¹ 1.73 m ⁻²) | UACR (mg g ⁻¹) | Rate of progression in the control arm (ml min ⁻¹ 1.73 m ⁻² /year) | Interventions evaluated |
|----------------|----------------------------------|-------------|-------------|--------------|--|----------------------------|--|--|
| All | 48 (85,681) | 62 (11) | 28,931 (34) | 75,194 (88) | 61.9 (46.9) | 257 (25–970) | -3.95 (2.10) | See below |
| Disease groups | | | | | | | | |
| CKD | 23 (12,335) | 54 (14) | 4637 (38) | 2,880 (23) | 45.6 (21.4) | 431 (60–1,240) | -4.01 (1.91) | RASB vs CCB, RASB vs control, low v usual BP, low vs usual diet, albuminuria targeted protocol, SGLT2I |
| Diabetes | 18 (72,309) | 64 (9) | 23,944 (33) | 72,309 (100) | 64.6 (46.2) | 213 (19–888) | -3.78 (2.16) | RASB vs CCB, RASB vs control, low vs usual BP, antiplatelet, SGLT2I, MRA, DPP4i, GLP-1RA, ERA |
| IgAN | 7 (1,037) | 40 (12) | 350 (34) | 5 (0) | 71.4 (30.2) | 1069 (719–1,557) | -3.26 (1.92) | Immunosuppression, RASB vs control |

Numbers in parentheses indicate the s.d. for continuous variables, percentage for dichotomous variables and 25th to 75th percentiles for UACR. BP, blood pressure; CCB, calcium channel blocker; DPP4i, dipeptidyl peptidase-4 inhibitor; MRA, mineralocorticoid receptor antagonist; RASB, renin-angiotensin system blocker; vs, versus.

across a broad collection of diverse clinical trials⁹. Using a smaller number of randomized controlled trials (RCTs), we previously demonstrated that treatment effects on albuminuria were associated with treatment effects on clinical kidney outcomes¹⁰. On the basis of this and other evidence, change in albuminuria is now used to evaluate drug efficacy in less common conditions such as IgA nephropathy (IgAN) or is used as a bridging biomarker for evidence translation of drug efficacy in one population to another^{11,12}. However, uncertainty persists about the use of change in albuminuria as a surrogate endpoint in more common causes of kidney disease and across severity of kidney diseases, limiting efficacy assessments of new therapies in many settings.

Since the publication of the earlier meta-analyses, several additional RCTs have been published, demonstrating that newer therapies reduce albuminuria and slow GFR decline, including sodium-glucose co-transporter-2 inhibitors (SGLT2Is), glucagon-like peptide-1 receptor agonist (GLP-1RA), endothelin receptor antagonists (ERAs) and a non-steroidal mineralocorticoid receptor antagonist (nsMRA)^{13–15}. These new therapies offer the opportunity to determine if the previously observed associations between treatment effects on albuminuria and treatment effects on clinical kidney outcomes persist across a wider range of interventions and disease settings. In addition, with inclusion of a larger number of studies, the consistency of the associations between treatment effects on albuminuria and kidney failure can be more robustly assessed across different stages of disease. Collectively, these data could further support the role of albuminuria as a surrogate endpoint for progression of CKD in future clinical trials.

Results

Study flow and characteristics

As previously described, RCTs of CKD progression were identified using a systematic literature search¹⁶. For this analysis, we updated our search and included seven new RCTs, leading to a total of 48 studies involving 12 treatment comparisons (Extended Data Fig. 1 and Supplementary Table 1). The baseline characteristics of the 85,681 included participants in 48 studies are summarized in Table 1, stratified by CKD etiologies, and in Supplementary Table 2. Bias assessment of included studies from the systematic review is presented in Supplementary Fig. 1. Across CKD etiologies, the mean age of the study participants ranged from 40 years to 64 years, and the proportion of women ranged from 33% to 38%. Mean baseline GFR ranged from 45.6 (s.d. 21) to 71.4 (s.d. 30) ml min⁻¹1.73 m⁻², and median baseline UACR ranged from 213 (25th to 75th percentiles, 19–888) mg g⁻¹ to 1,069 (25th to 75th percentiles, 719–1,557) mg g⁻¹.

Treatment effects on UACR change and the clinical endpoint

Over a 6-month (25th to 75th percentiles, 6.0–6.4) period, the overall geometric mean ratio (GMR) from baseline in UACR in the control and active treatment arms was 0.84 (95% confidence interval: 0.78–0.89) and 0.64 (95% confidence interval: 0.59–0.68), respectively, corresponding to a pooled average 25% reduction from baseline in UACR due to the treatment (GMR 0.75 (95% confidence interval: 0.71–0.79); Extended Data Fig. 2). Results over 12 months and including two studies with tabular data were similar (Extended Data Fig. 2 and Supplementary Fig. 2). Overall, the active interventions reduced the rate of the clinical endpoint compared to the control by 24% (hazard ratio (HR) = 0.76 (95% confidence interval: 0.71–0.82), with no clear evidence that results differ across subgroups; Extended Data Fig. 2). The early treatment effects on UACR were similar by the type or severity of kidney disease (Extended Data Fig. 2).

Trial-level analysis in the overall study population

Results of the trial-level analysis evaluating the association between the treatment effects on 6-month UACR change to those of the clinical endpoint are reported in Fig. 1. The intercept of the meta-regression line was -0.05 (95% BCI: -0.22 to 0.11). An intercept with BCI that crosses zero indicates that, with no reduction in UACR in the active compared to control treatment, there was a low probability of having a substantial average treatment effect on the clinical outcome. The meta-regression slope relating treatment effects on the log scale was 0.68 (95% BCI: 0.17–1.19), which indicates that every 30% greater reduction in the geometric mean UACR was associated with an average 19% (95% BCI: 5–30%) lower hazard for the clinical endpoint. The Bayesian approach used in our analysis can be used to provide estimates of the probabilities that the meta-regression coefficients and R^2 values, which characterize the relationships between the treatment effects on UACR and the clinical endpoint, fall into different ranges (Fig. 2). The posterior probability that the meta-regression slope was positive, greater than zero, was 99.5%. The posterior medians for R^2 and residual mean square error (RMSE) were 0.66 (95% BCI: 0.06–0.98) and 0.08 (95% BCI: 0.02–0.17), respectively. The wide BCIs decrease our certainty about the strength of the association. Thus, although the trial-level association showed a 45.2% chance of a strong surrogate endpoint (as defined as $R^2 > 0.72$), there are 28.0% and 26.8% probabilities for the R^2 falling into the low ($R^2 < 0.49$) and moderate (R^2 , 0.49–0.72) ranges, respectively, for the strength of a surrogate endpoint. Despite the wide 95% BCI for the R^2 , for an observed 30% reduction in UACR the posterior probability of a beneficial treatment effect on the clinical endpoint was greater than 99%. Results were consistent in sensitivity analysis using the alternative clinical endpoint of kidney failure alone (Extended Data Table 1 and

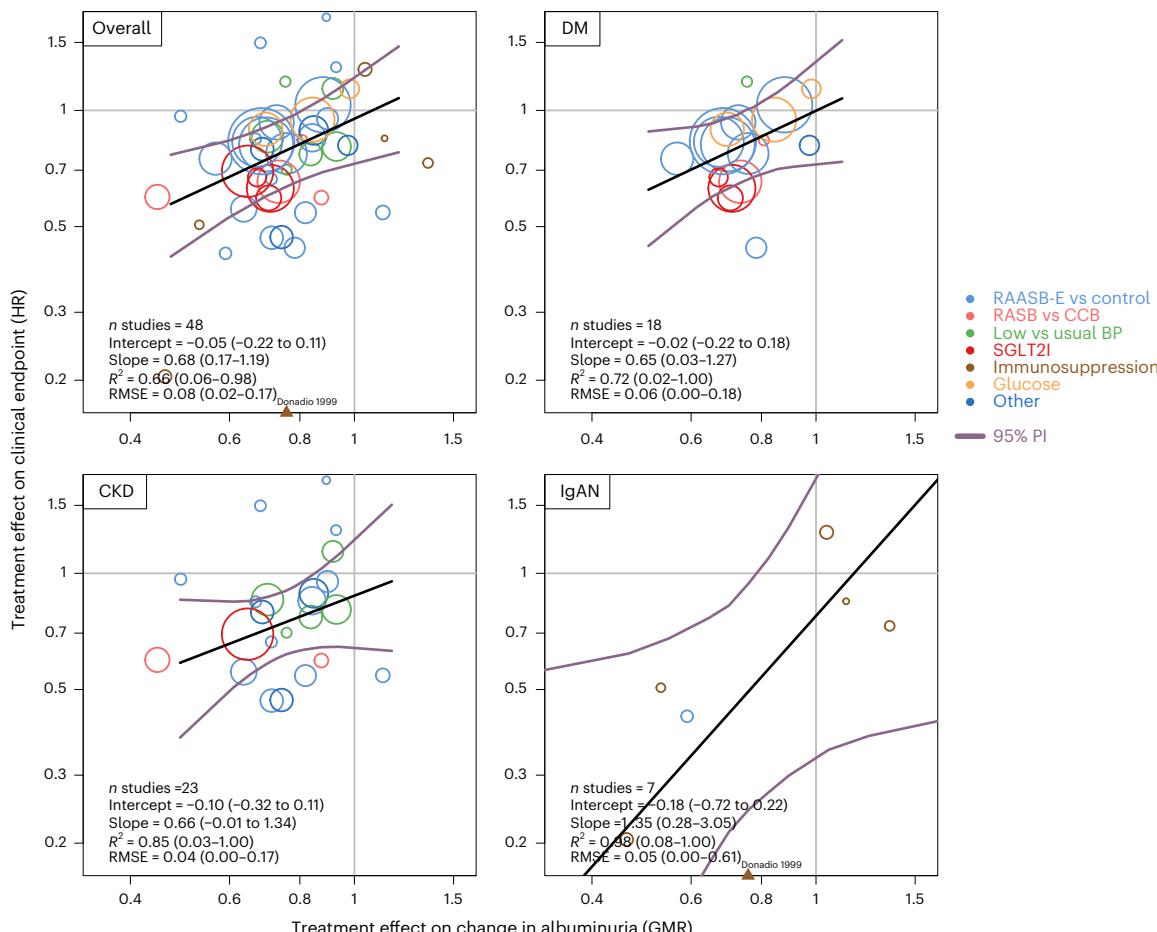


Fig. 1 | Trial-level association between treatment effects on UACR and treatment effects on the clinical endpoint, overall and by disease subgroup.

The treatment effects on UACR are expressed on the horizontal axis.

The treatment effects on the clinical endpoint are expressed on the vertical axis.

Treatment effects on UACR are expressed as GMRs. Each circle

represents a separate study, with the size of the circle proportional to the number of events. The colors of the circles indicate intervention type. The black line indicates the regression line. The plum-colored line indicates the Bayesian 95% prediction interval (PI). BP, blood pressure; CCB, calcium channel blocker; DM, diabetes mellitus.

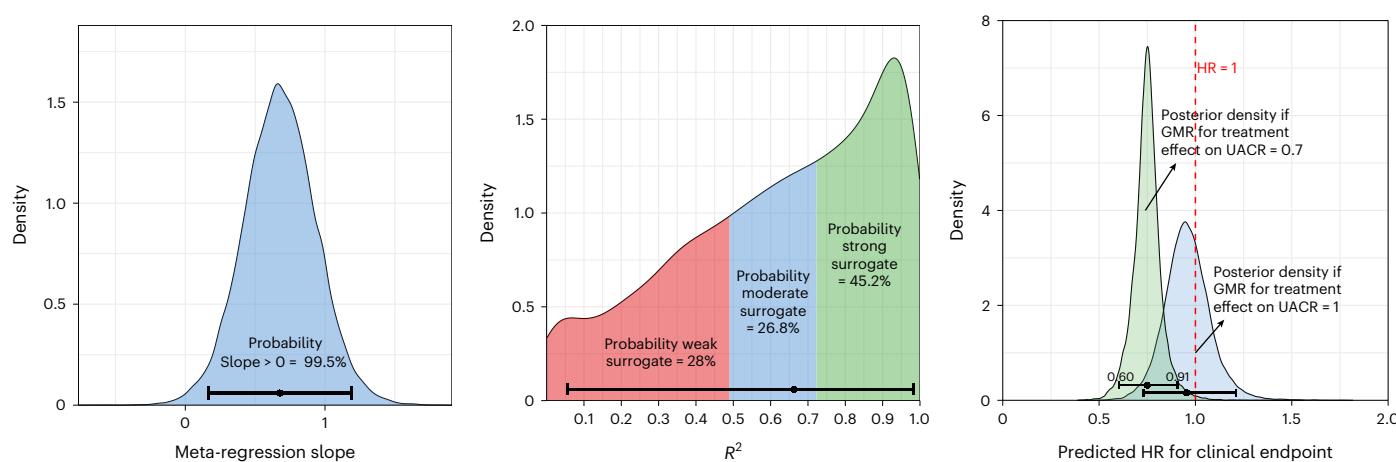


Fig. 2 | Posterior distributions for three aspects of the trial-level association between the treatment effects on UACR at 6 months and the clinical endpoint.

The graphs show the full posterior distributions of the meta-regression slope (left panel) and R² (middle panel) and the posterior predictive distribution of true treatment effects on the clinical endpoint given selected true treatment effects on UACR (right panel). The black intervals in the lower part of each panel extend from the 2.5th to the 97.5th percentiles of the relevant posterior distribution.

The numbers above the line in the right panel are the values of the 2.5th and 97.5th percentiles for the posterior distribution when the treatment effect on UACR is given by a GMR of 0.7. More than 99% of the area under the green curve in the right panel is below 1, indicating that the posterior probability of beneficial treatment effects on the clinical endpoint for an observed 30% reduction in UACR was greater than 99%.

Table 2 | Trial-level analysis overall and at designated levels of CKD severity measures

| Study classifier | Severity level | Meta-regression slope | Intercept | R ² | RMSE |
|--|----------------|-----------------------|-----------------------|------------------|------------------|
| Overall | | 0.68 (0.17–1.19) | -0.05 (-0.22 to 0.11) | 0.66 (0.06–0.98) | 0.08 (0.02–0.17) |
| Baseline GFR (ml min ⁻¹ 1.73 m ⁻²) ^a | 45 | 0.58 (0.07–1.09) | -0.08 (-0.25 to 0.09) | 0.67 (0.03–0.98) | 0.06 (0.02–0.17) |
| | 75 | 1.57 (0.56–2.65) | 0.25 (-0.08 to 0.59) | 0.94 (0.43–1.00) | 0.06 (0.02–0.17) |
| | 100 | 0.87 (-0.05 to 1.87) | 0.12 (-0.18 to 0.44) | 0.77 (0.02–0.99) | 0.07 (0.02–0.17) |
| Baseline UACR (mg g ⁻¹) ^a | 300 | 0.68 (0.16–1.23) | -0.01 (-0.18 to 0.17) | 0.68 (0.06–0.98) | 0.07 (0.02–0.17) |
| | 1,000 | 0.47 (-0.27 to 1.24) | -0.15 (-0.40 to 0.09) | 0.52 (0.00–0.97) | 0.07 (0.02–0.17) |
| | -5 | 0.76 (0.02–1.54) | -0.02 (-0.26 to 0.21) | 0.67 (0.02–0.99) | 0.09 (0.02–0.19) |
| GFR rate of progression (ml min ⁻¹ 1.73 m ⁻² /year) ^{a,b} | -3 | 0.64 (-0.05 to 1.32) | -0.07 (-0.29 to 0.14) | 0.58 (0.01–0.98) | 0.09 (0.02–0.19) |
| | -1 | 0.52 (-0.97 to 1.97) | -0.11 (-0.54 to 0.30) | 0.58 (0.00–0.99) | 0.09 (0.02–0.19) |

Posterior medians and 95% BCIs are shown. The meta-regression slope, intercept, R² and RMSE are provided for the overall analysis (top row) and in subsequent rows under extended models with CKD severity interaction terms at designated levels of baseline GFR, baseline UACR and the GFR rate of progression in the control group. The meta-regression slope indicates the change in the log-transformed HR for the clinical endpoint associated with each 1-unit increase in the treatment effect on log-transformed UACR. See Fig. 2 for expression of the meta-regression in terms of the HR itself. ^aThe differences in posterior median meta-regression slope displayed in the table are consistent with chance variation for baseline UACR and the GFR rate of progression but exceed the expected chance variation for baseline GFR. See also Extended Data Table 2. ^bGFR progression was defined by the mean GFR chronic slope in the control arm of the trial.

Extended Data Fig. 3). Results also remained similar after inclusion of tabular data from two new studies^{17,18} that assessed the effect of a SGLT2I and GLP-1RA (Supplementary Fig. 3). Results were also consistent for sex-specific subgroups (Supplementary Table 3).

Trial-level analysis by disease and intervention classes

The posterior median meta-regression slope describing the associations between treatment effects on UACR change and those of the clinical endpoint appeared to be greater for IgAN studies compared to diabetes or CKD studies of unspecified cause, although the BCIs around the slopes for the disease-specific meta-regressions were very wide, precluding a definitive conclusion of a difference (Fig. 1).

There was evidence of heterogeneity in the association of treatment effects on UACR and those of the clinical endpoint as a function of baseline level of GFR. The 95% BCI around the interaction term excluded zero (Table 2, Extended Data Table 2 and Extended Data Fig. 4a). Specifically, the meta-regression slope appeared to be larger in magnitude, and the meta-regression intercept increasingly positive, for trials with higher baseline GFR (Table 2, Extended Data Table 2 and Extended Data Fig. 4a). The association of the meta-regression slope with baseline GFR weakened but remained present after exclusion of the IgAN studies, which had higher baseline GFR compared to the other studies (Extended Data Fig. 4b). There was no clear evidence that the meta-regression slope differed across different levels of baseline UACR or mean GFR progression rates.

When the meta-regression model was extended to allow separate meta-regressions within subgroups defined by intervention classes, we found that, for renin–angiotensin–aldosterone system blockade or endothelin receptor antagonists (RAASB-E), an estimated 82% of the average effect of the intervention was accounted for by the average effect of the intervention on UACR (Table 3). The treatment effects on UACR accounted for smaller percentages of the average total treatment effect on the clinical endpoint for renin–angiotensin system (RAS) inhibitors versus calcium channel blockers, SGLT2Is and immunosuppression subgroups, suggesting that the estimated treatment benefits may have exceeded the benefits predicted from the average UACR effects of these intervention classes. Although the wide BCIs for these subgroup analyses preclude definitive conclusions, these results are consistent with the possibility that beneficial effects of these treatment classes on the clinical endpoint may not be fully explained by the effects of the respective interventions on UACR alone.

Prediction intervals and positive predictive value

For application of UACR as a surrogate endpoint in future RCTs, Table 4, Fig. 2, Extended Data Table 3 and Extended Data Fig. 5 show

the predicted HRs and 95% prediction intervals for the treatment effects on the clinical endpoint calculated across a range of possible UACR treatment effects. For a study of sample size of 200 patients, treatments that reduce the geometric mean albuminuria by 30% versus the control group will provide an average HR for the clinical endpoint of 0.75, with 95% probability that the true HR is between 0.57 and 0.95. Table 4, Fig. 2 and Extended Data Table 3 also show that, with the same sample size, a treatment effect corresponding to a 27% reduction in UACR would be required to provide a 97.5% probability of clinical benefit defined by HR < 1. Similar results are seen for changes in albuminuria at 12 months (Extended Data Table 3). Larger UACR reductions are required to infer clinical benefit defined by HR < 0.8 (Extended Data Table 4).

Discussion

Demonstrating the validity of surrogate endpoints requires evaluation of causal pathways of the surrogate endpoint in the disease process; epidemiological data demonstrating a strong and consistent association between the surrogate and clinical endpoints; and trial-level analyses, as presented here, demonstrating a consistent association between treatment effects on the surrogate and treatment effects on the clinical endpoint. Our meta-analysis including 48 RCTs and comprising more than 85,000 participants provides, to our knowledge, the most comprehensive analysis to date and demonstrates that treatment effects on UACR at 6 months accounted for an estimated 66% and 74% of the variation between studies in the treatment effects on the composite clinical kidney endpoint and kidney failure, respectively. Although these trial-level R² estimates correspond to a moderately strong surrogate, there was substantial uncertainty reflected by the wide BCIs. Despite the uncertainty in the R² value, the analyses also demonstrate that treatments that reduce UACR by at least 25% have a high probability of reducing the risk for the clinical endpoint, whereas treatments that do not reduce albuminuria will typically not substantially benefit the clinical endpoint. These results may, therefore, aid in decision-making for regulatory drug approval in settings where albuminuria can be used as a conditional endpoint or inform design of phase 3 trials. Collectively, these results using an updated and expanded dataset of clinical trials with new interventions that reduce albuminuria and the risk of kidney failure confirm that change in albuminuria can be considered a reasonably likely surrogate endpoint across various diseases and patient subgroups¹⁰.

These analyses extend the results of our previous meta-analysis across more heterogeneous interventions, study populations and disease severities. Specifically, our previous analysis mainly included studies of RAS inhibitors, whereas new drugs for the treatment of CKD have now been approved, including SGLT2I, an nsMRA, a GLP-1RA and

Table 3 | Proportion of treatment effect on the clinical endpoint explained by UACR for different intervention classes

| Interventions | Average predicted overall hazard reduction for the clinical endpoint | Average predicted hazard reduction for the clinical endpoint after controlling for the treatment effect on UACR | Percent of expected overall hazard reduction explained by the average treatment effect on UACR |
|-------------------|--|---|--|
| RAASB-E | 0.19 (0.11–0.27) | 0.04 (−0.21 to 0.25) | 82% |
| RASB vs CCB | 0.33 (0.14–0.47) | 0.24 (−0.15 to 0.53) | 34% |
| Low vs usual BP | 0.14 (−0.05 to 0.31) ^a | 0.04 (−0.38 to 0.31) | NA ^a |
| Immunosuppressant | 0.50 (0.18–0.70) | 0.35 (−0.17 to 0.62) | 41% |
| SGLT2I | 0.34 (0.21–0.45) | 0.21 (−0.43 to 0.63) | 45% |
| Intensive glucose | 0.06 (−0.17 to 0.25) ^a | −0.07 (−0.50 to 0.21) | NA ^a |
| Other | 0.27 (0.06–0.45) | 0.16 (−0.22 to 0.42) | 48% |

Posterior medians and 95% BCIs are shown. Data in the final column show ratios of posterior medians from the preceding two columns expressed as percentages. ^a The proportion of the treatment effect explained by the early treatment effect on UACR was not computed when the 95% BCI for the HR of the overall effect of treatment class included 1. BP, blood pressure; CCB, calcium channel blocker; NA, not applicable; RASB, renin–angiotensin system blocker; vs, versus.

Table 4 | Application to a future trial: predicted treatment effects on the clinical endpoint for varying percent reductions in UACR

| Observed treatment effect on change in UACR | RCT of 1,600 | | RCT of 800 | | RCT of 200 | |
|--|---------------------------------------|----------------------|---------------------------------------|----------------------|---------------------------------------|----------------------|
| | Median HR and 95% prediction interval | PPV _{trial} | Median HR and 95% prediction interval | PPV _{trial} | Median HR and 95% prediction interval | PPV _{trial} |
| 0.5 | 0.60 (0.44–0.78) | 1.00 | 0.60 (0.44–0.78) | 1.00 | 0.60 (0.42–0.79) | 1.00 |
| 0.6 | 0.67 (0.53–0.84) | 1.00 | 0.68 (0.52–0.84) | 1.00 | 0.68 (0.50–0.86) | 1.00 |
| 0.7 | 0.75 (0.60–0.91) | 0.99 | 0.75 (0.60–0.92) | 0.99 | 0.75 (0.57–0.95) | 0.99 |
| 0.8 | 0.82 (0.65–1.00) | 0.97 | 0.82 (0.65–1.01) | 0.97 | 0.82 (0.63–1.05) | 0.95 |
| 0.9 | 0.89 (0.69–1.11) | 0.88 | 0.89 (0.69–1.11) | 0.87 | 0.88 (0.68–1.17) | 0.84 |
| 1.0 | 0.95 (0.73–1.22) | 0.67 | 0.95 (0.73–1.23) | 0.66 | 0.94 (0.72–1.29) | 0.66 |
| Threshold to assure PPV _{trial} ≥ 97.5% | 0.79 | 0.78 | | | 0.73 | |

Calculations are shown for clinical trials with 1,600, 800 or 200 participants. All calculations assume an s.d. of the change in log UACR of 0.75 and that the future trial allocates equal numbers of patients to a single active treatment and a single control group. Treatment effects on UACR are expressed as the GMRs of follow-up UACR in the active treatment arm compared to the control arm. PPV_{trial} represents the probability that the HR for the clinical endpoint is less than 1. See Extended Data Table 4 for probabilities that the HR for the clinical endpoint is less than 0.8.

an ERA. Inclusion of these new drug classes enabled us to assess the performance of albuminuria as a surrogate endpoint across a wider range of therapies. Moreover, the larger population with different types of CKD and at various stages of disease increased the generalizability of the trial-level associations observed and strengthened the evidence for the validity of albuminuria change as a surrogate endpoint for kidney failure.

Detecting trial-level associations requires heterogeneity in study populations to have sufficient power. However, such heterogeneity may raise the question as to whether the overall results can be applied to specific subgroups. We observed wide BCIs for the R^2 metric, which adds further emphasis to this question. In this updated meta-analysis, we used more rigorous methods to examine whether the trial-level associations held by kidney disease etiology, CKD disease severity and key interventions. The median trial-level R^2 was higher in each kidney-disease-defined subgroup compared to the meta-regression R^2 across the full set of studies. This may be explained by the fact that our partial pooling model allowed the meta-regression parameters to vary, enabling the subgroup-specific meta-regressions to better fit the data within each disease-defined subgroup relative to the overall model across the full cluster of trials. Nevertheless, the smaller number of trials within each intervention class, and the fact that variation in treatment effects on albuminuria is limited within the same treatment class, results in uncertainty in the corresponding credible intervals. We also observed stronger associations between change in albuminuria and clinical outcomes for studies with higher baseline GFR. We think that these results are possibly owing to the different clinical trial

populations and trial inclusion criteria with different baseline GFR levels and rates of progression. For example, in sensitivity analyses after removing IgAN studies, which had higher levels of baseline GFR than the other studies, the differential results by level of GFR were attenuated.

We previously demonstrated the association between early treatment effects on albuminuria and treatment effects on clinical endpoint and GFR slope in IgAN. The current data are an advance on this work, as we now demonstrate the associations on clinical endpoints and use a partial pooling model to improve precision in subgroup-specific meta-regression parameter estimates. Despite that this model induces information sharing across subgroups, the meta-regression slope remained strongest within the IgAN studies. These data, along with other meta-analyses, support the concept that albuminuria can be considered a reasonably likely surrogate endpoint for conditional approval for new therapies in IgAN¹¹.

Albuminuria can be applied as a surrogate endpoint in future trials for interventions in which reducing albuminuria is hypothesized to be one of the main kidney-protective mechanisms. It is, therefore, relevant to understand to what extent kidney-protective effects of different drug classes can be explained by the reduction in albuminuria. Our results suggest that the kidney-protective effects of SGLT2I and immunosuppression are partly explained by the early reduction in albuminuria, whereas the effects of RAAS inhibitors can be attributed to a large extent to the early reduction in albuminuria. These results complement previous observational studies reporting that, for RAAS inhibitors, more than 90% of the protective effect is explained by the reduction in albuminuria, whereas, for SGLT2I, this is approximately

40%^{19–22}. The advantage of the present trial-level analysis over observational individual patient-level analysis is that the former specifically relates treatment effects on albuminuria to treatment effects on long-term kidney outcomes and avoids the risk of confounding between early albuminuria changes and other patient factors. On the other hand, the relatively small numbers of trials within individual drug classes and wide prediction intervals for the hazard reductions associated with the treatment before and after controlling for the treatment effects on albuminuria limit the precision of the current analysis.

The present study accompanies a previous meta-analysis of RCTs on the validity of the rate of GFR decline (GFR slope) as a surrogate endpoint for kidney failure. Both albuminuria and GFR slope have their strengths and weaknesses. GFR slope has a greater predictive ability for kidney failure than albuminuria. In addition, albuminuria can be used only as a surrogate endpoint for populations with some level of albuminuria and for interventions whose mechanism of action involve albuminuria-mediated pathophysiological pathways. Analysis of GFR slope is more complex than change in albuminuria. Specifically, GFR slope requires a longer follow-up for many clinical trial settings, and the frequently encountered initial effect of the intervention on GFR, which differs from the long-term effect of the intervention, creates complexities in the interpretation of the overall effect of the intervention on GFR slope²³. The complementary strengths and limitations of both surrogates enabled us to develop methods that use the combination of treatment effects on both albuminuria and GFR²⁴. This joint endpoint can be used when there is uncertainty whether the mechanism of action of the intervention involves albuminuria-mediated pathways—for example, in the case of anti-inflammatory or anti-fibrotic interventions.

These results have several implications. Treatment effects on change in albuminuria are used in phase 2 studies to determine drug efficacy and the optimal dose for confirmatory outcome trials, in phase 3 trials in rare kidney diseases for conditional regulatory approval or when treatments with established efficacy and safety in one population are expanded to another population. For example, the FINE-ONE trial uses change in albuminuria as an endpoint for potential label expansion of finerenone for slowing the progression of GFR decline to adults with type 1 diabetes and CKD²⁵. The requirement for changes in albuminuria of approximately 25% to have a high confidence that the drug will exert clinical benefit provides guidance to drug developers, regulators and clinicians to determine drug efficacy. Various new interventions, including ERAs, aldosterone synthase inhibitors, incretin analogues and soluble guanyl cyclase inhibitors, have shown reductions in albuminuria of 25% or more in phase 2 clinical trials^{26–29}. Their effects on clinical kidney outcomes are now further investigated in confirmatory phase 3 clinical trials. These trials will also establish the safety of these interventions, which cannot be adequately studied in shorter phase 2 studies, with change in albuminuria as a surrogate endpoint. The overall drug development program would consider how to best determine drug safety and long-term efficacy.

A key strength of this study is the availability of individual patient-level data, which enabled harmonization of endpoint definitions and detailed subgroup analyses. A limitation is that we were not able to include all relevant clinical trials due to data-sharing restrictions, which may have affected generalizability and precision of effect estimates, in particular for subgroup analyses. The precision of the effect estimates may also be affected by heterogeneity in the study designs and procedures to assess albuminuria. A few recently completed trials in IgAN could not be included because there were insufficient clinical endpoints for inclusion in our study. We also note that, although sample sizes of studies in patients with diabetes were larger compared to CKD or IgAN studies, the number of studies, which is the unit of measurement in a trial-level analysis, did not differ between diabetes and CKD. Second, although there was uncertainty in the estimate of the intercept of the trial-level association, the small but negative intercept of −0.05 means that the median predicted hazard

reduction for the clinical endpoint is approximately 5% in favor of the treatment even in the absence of a treatment effect on albuminuria. We also note that these results apply to interventions whose mechanism of action involve albuminuria-mediated pathophysiological pathways. Similarly, these results apply to kidney diseases with at least minimal albuminuria at baseline ($ACR > 30 \text{ mg g}^{-1}$) and that were characterized by albuminuria-mediated pathophysiological mechanisms. We do not recommend use of albuminuria as a surrogate endpoint in kidney diseases characterized by low albuminuria. Finally, the included studies typically enrolled a lower proportion of women than men. However, results from the additional analysis by sex showed similar results as our main analysis. The included studies also enrolled few young adults, and the number of participants with advanced CKD stage 3b or 4 was low. Therefore, the statistical power was insufficient to comprehensively assess the performance of albuminuria as a surrogate endpoint in these populations.

In summary, the results of this updated meta-analysis support use of early changes in albuminuria as a reasonably likely surrogate for kidney failure in clinical trials enrolling participants with moderate or severe albuminuria.

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/s41591-025-04057-z>.

References

1. Jager, K. J. et al. A single number for advocacy and communication-worldwide more than 850 million individuals have kidney diseases. *Kidney Int.* **96**, 1048–1050 (2019).
2. Kidney Disease: Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2024 Clinical practice guideline for the evaluation and management of chronic kidney disease. *Kidney Int.* **105**, S117–S314 (2024).
3. Inker, L. A. et al. Clinical trial considerations in developing treatments for early stages of common, chronic kidney diseases: a scientific workshop cosponsored by the National Kidney Foundation and the US Food and Drug Administration. *Am. J. Kidney Dis.* **80**, 513–526 (2022).
4. FDA-NIH Biomarker Working Group. BEST (Biomarkers, EndpointS, and other Tools) Resource. <https://www.ncbi.nlm.nih.gov/books/NBK338448/> (National Institutes of Health, 2016).
5. Tang, S. et al. Albumin stimulates interleukin-8 expression in proximal tubular epithelial cells in vitro and in vivo. *J. Clin. Invest.* **111**, 515–527 (2003).
6. Morigli, M. et al. Protein overload-induced NF-κB activation in proximal tubular cells requires H₂O₂ through a PKC-dependent pathway. *J. Am. Soc. Nephrol.* **13**, 1179–1189 (2002).
7. Faivre, A. et al. Spatiotemporal landscape of kidney tubular responses to glomerular proteinuria. *J. Am. Soc. Nephrol.* **35**, 854–869 (2024).
8. Coresh, J. et al. Change in albuminuria and subsequent risk of end-stage kidney disease: an individual participant-level consortium meta-analysis of observational studies. *Lancet Diabetes Endocrinol.* **7**, 115–127 (2019).
9. Joffe, M. M. & Greene, T. Related causal frameworks for surrogate outcomes. *Biometrics* **65**, 530–538 (2009).
10. Heerspink, H. J. L. et al. Change in albuminuria as a surrogate endpoint for progression of kidney disease: a meta-analysis of treatment effects in randomised clinical trials. *Lancet Diabetes Endocrinol.* **7**, 128–139 (2019).
11. Thompson, A. et al. Proteinuria reduction as a surrogate end point in trials of IgA nephropathy. *Clin. J. Am. Soc. Nephrol.* **14**, 469–481 (2019).

12. International Society of Glomerular Disease. PARASOL: proteinuria as clinical trial endpoints in focal segmental glomerulosclerosis <https://www.is-gd.org/parasol> (accessed 9 October 2025).
13. Perkovic, V. et al. Canagliflozin and renal outcomes in type 2 diabetes and nephropathy. *N. Engl. J. Med.* **380**, 2295–2306 (2019).
14. Heerspink, H. J. L. et al. Dapagliflozin in patients with chronic kidney disease. *N. Engl. J. Med.* **383**, 1436–1446 (2020).
15. Heerspink, H. J. L. et al. Atrasentan and renal events in patients with type 2 diabetes and chronic kidney disease (SONAR): a double-blind, randomised, placebo-controlled trial. *Lancet* **393**, 1937–1947 (2019).
16. Inker, L. A. et al. A meta-analysis of GFR slope as a surrogate endpoint for kidney failure. *Nat. Med.* **29**, 1867–1876 (2023).
17. Perkovic, V. et al. Effects of semaglutide on chronic kidney disease in patients with type 2 diabetes. *N. Engl. J. Med.* **391**, 109–121 (2024).
18. The EMPA-KIDNEY Collaborative Group. Empagliflozin in patients with chronic kidney disease. *N. Engl. J. Med.* **388**, 117–127 (2023).
19. Li, J. et al. Mediators of the effects of canagliflozin on kidney protection in patients with type 2 diabetes. *Kidney Int.* **98**, 769–777 (2020).
20. Oshima, M. et al. Early change in albuminuria with canagliflozin predicts kidney and cardiovascular outcomes: a post hoc analysis from the CREDENCE trial. *J. Am. Soc. Nephrol.* **31**, 2925–2936 (2020).
21. Agarwal, R. et al. Impact of finerenone-induced albuminuria reduction on chronic kidney disease outcomes in type 2 diabetes: a mediation analysis. *Ann. Intern. Med.* **176**, 1606–1616 (2023).
22. de Zeeuw, D. et al. Proteinuria, a target for renoprotection in patients with type 2 diabetic nephropathy: lessons from RENAAL. *Kidney Int.* **65**, 2309–2320 (2004).
23. Heerspink, H. J. L. et al. A meta-analysis of randomized controlled clinical trials for implications of acute treatment effects on glomerular filtration rate for long-term kidney protection. *Kidney Int.* **106**, 688–698 (2024).
24. Heerspink, H. J. L. et al. Change in albuminuria and GFR Slope as joint surrogate end points for kidney failure: implications for phase 2 clinical trials in CKD. *J. Am. Soc. Nephrol.* **34**, 955–968 (2023).
25. Heerspink, H. J. L. et al. Rationale and design of a randomised phase III registration trial investigating finerenone in participants with type 1 diabetes and chronic kidney disease: the FINE-ONE trial. *Diabetes Res. Clin. Pract.* **204**, 110908 (2023).
26. Apperloo, E. M. et al. Semaglutide in patients with overweight or obesity and chronic kidney disease without diabetes: a randomized double-blind placebo-controlled clinical trial. *Nat. Med.* **31**, 278–285 (2024).
27. Heerspink, H. J. L. et al. Effect of avemacicoustat on albuminuria in patients with CKD: two randomized placebo-controlled trials. *J. Am. Soc. Nephrol.* **35**, 1227–1239 (2024).
28. Heerspink, H. J. L. et al. The selective endothelin receptor antagonist SC0062 in IgA nephropathy: a randomized double-blind placebo-controlled clinical trial. *J. Am. Soc. Nephrol.* **36**, 657–667 (2025).
29. Tuttle, K. R. et al. Efficacy and safety of aldosterone synthase inhibition with and without empagliflozin for chronic kidney disease: a randomised, controlled, phase 2 trial. *Lancet* **403**, 379–390 (2024).

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2025

¹Department of Clinical Pharmacy and Pharmacology, University of Groningen, University Medical Center Groningen, Groningen, Netherlands. ²George Institute for Global Health, University of New South Wales, Sydney, New South Wales, Australia. ³Population Health Sciences, University of Utah School of Medicine, Salt Lake City, UT, USA. ⁴Department of Population and Public Health Sciences, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA. ⁵Division of Nephrology, Tufts Medical Center, Boston, MA, USA. ⁶Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, MA, USA. ⁷Tufts Clinical and Translational Science Institute, Tufts University, Boston, MA, USA. ⁸Division of Nephrology, Columbia University Medical Center and the New York Presbyterian Hospital, New York, NY, USA. ⁹Instituto de Investigación Hospital 12 de octubre (i+12), Madrid, Spain. ¹⁰Division of Nephrology, RWTH Aachen University, Aachen, Germany. ¹¹University of Strasbourg, AURAL, Strasbourg, France. ¹²Nakayamadera Imai Clinic, Takarazuka, Japan. ¹³Program in Health Services and Systems Research, Duke-NUS Medical School, Singapore, Singapore. ¹⁴Division of Nephrology, Vanderbilt University Medical Center, Nashville, TN, USA. ¹⁵Division of Nephrology, Prince of Wales Hospital, Chinese University of Hong Kong, Shatin, Hong Kong. ¹⁶Department of Nephrology, Alessandro Manzoni Hospital, ASST Lecco, Lecco, Italy. ¹⁷Department of Nephrology, AZ Delta, Roeselare, Belgium. ¹⁸Istituto di Ricerche Farmacologiche Mario Negri IRCCS, Bergamo, Italy. ¹⁹Renal, Dialysis and Transplant Unit, Department of Emergency and Organ Transplantation, University of Bari, Bari, Italy. ²⁰Division of Nephrology, University Hospital of Würzburg, Würzburg, Germany.

✉ e-mail: h.j.lambers.heerspink@umcg.nl

Methods

Datasets and analytical groups

As previously described, RCTs of CKD progression were identified using a systematic literature search¹⁶. For this study, we updated our search and included seven additional treatment comparisons (herein referred to as studies), resulting in a total of 73 potential eligible studies (Supplementary Table 1, Supplementary Fig. 1 and Supplementary Notes 1 and 2). We excluded studies without baseline or follow-up albuminuria assessments ($n = 18$), studies with interventions that were hypothesized not to affect albuminuria ($n = 2$) and studies in disease areas that could not be pooled with another study for subgroup analysis ($n = 1$). We excluded glomerulonephritis studies other than IgAN given the goal of examining results in subgroups by CKD etiology and the lack of a sufficient number of studies in other glomerular disease subgroups ($n = 3$). This led to the inclusion of 48 studies comprising 12 treatment comparisons. In a sensitivity analysis, we included two recent studies for which we received from the clinical trial investigators only tabular data but not individual participant data^{17,18}.

Change in albuminuria

The measures of albuminuria varied across studies, with most measuring protein excretion rate. Because the Kidney Disease: Improving Global Outcomes (KDIGO) guidelines for the management of CKD recommend use of UACR, we expressed all urine protein measures to UACR using a commonly used conversion: UACR mg g^{-1} (mg mmol^{-1}) = $0.6 \times \text{protein excretion rate (mg d}^{-1}\text{)}$ ³⁰. For further reporting of methods and results, we use the term UACR. An early change in UACR was defined as the change in log-transformed UACR from baseline to the measurement closest to 6 months (within 2.5 months and 14 months) or 12 months (within 2.5 and 19 months).

Clinical endpoints

The clinical endpoint was defined as a composite of kidney failure (defined as GFR < 15 ml min⁻¹ 1.73 m⁻², initiation of chronic treatment with dialysis or kidney transplantation) or sustained doubling of serum creatinine. GFR was calculated in all trials using the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) 2009 equation³¹. We used the alternative clinical endpoint of kidney failure alone as our endpoint in a sensitivity analysis.

Statistical analyses

Objectives. The statistical analyses had four main objectives: (1) to determine the associations between early treatment effects on albuminuria and treatment effects on clinical endpoints across RCTs using trial-level Bayesian meta-regression analysis; (2) to evaluate the consistency of these associations across disease subgroups and across levels of trial-level baseline GFR, UACR and GFR rate of decline; (3) to estimate the extent to which the treatment effects on the clinical endpoint can be accounted for by the early treatment effects on UACR for different types of interventions; and (4) to use the results from the trial-level analysis to inform the application of UACR change as a surrogate endpoint in future RCTs by estimating the probability of clinical benefit associated with a range of treatment effects on UACR change.

Trial-level analysis. The trial-level analysis consisted of two stages: (1) intent-to-treat estimation of the treatment effects on the early change in UACR and on the clinical endpoint within each study using individual patient-level data and (2) Bayesian trial-level meta-regression to relate the treatment effects on the early change in UACR and the clinical endpoint across RCTs. In the first stage, treatment effects on UACR change were estimated by performing analyses of covariance within each study, with log UACR change from baseline as the endpoint, adjusting for treatment and log baseline UACR. Treatment effects on UACR were expressed as GMRs. Treatment effects on the clinical endpoint were estimated by Cox proportional hazard regressions to estimate

HRs for the treatment in each study. In the second stage, a Bayesian mixed-effects meta-regression analysis was performed to relate the estimated treatment effects on the early change in UACR with the estimated treatment effects on the clinical endpoint with study as the unit of analysis. The model associated the treatment effects on the two endpoints after accounting for random errors in the estimated effects and correlations in random errors within each RCT³². The meta-regression supports validity of UACR change as a surrogate endpoint if (1) the slope of the meta-regression line has a large magnitude with its 95% BCIs excluding zero; (2) the intercept is close to zero, implying absence of a substantial average effect on the clinical endpoint when the treatment does not affect ACR; (3) the R^2 is high, so that treatment effects on ACR account for a high percentage of the variation in treatment effects on the clinical endpoint; and (4) the RMSE is low, assuring low variation in the treatment effects on the clinical endpoint given a fixed treatment effect on ACR. We used the designations of low, moderate and strong trial-level association as defined by $R^2 < 0.49$, $0.49–0.72$ and ≥ 0.72 , respectively³³.

Analysis by type of kidney disease and severity of CKD. For subgroup analysis, the Bayesian hierarchical model used for the primary analysis was extended to allow for subgroup-specific meta-regression parameters. We used a partial pooling model, which facilitates a data-adaptive information sharing across subgroups to mitigate excessive uncertainty in parameter estimates due to small subgroup sizes. The within-subgroup meta-regression slope, intercept and RMSE are expected to be estimated with improved precision due to partial pooling, but separation can still be expected if the data indicate strong heterogeneity across subgroups³⁴.

We assessed heterogeneity in the trial-level association of interest across three measures of CKD severity: mean baseline GFR, geometric mean baseline UACR and the mean rate of disease progression (the control arm GFR chronic slope). For each severity measure, we added interaction terms between model coefficients and the severity measure to the primary meta-regression model to determine if the meta-regression intercept and slope varied as a function of trial-level disease severity³⁵.

Portion of treatment effects explained by early changes in albuminuria. We used a modified partial pooling model with subgroups defined by the treatment classes to evaluate the portion of the average treatment effect on the clinical endpoint accounted for by the effect on UACR for each treatment class. In this application of the partial pooling model, the intercepts of each class were estimated independently, without borrowing information from other treatment classes, whereas the meta-regression slopes and RMSEs were partially pooled across subgroups. We used the model to estimate the average hazard reduction of each treatment class on the clinical endpoint, with and without controlling for the average effect of the treatment on UACR in the same class, and computed the ratios of the average predicted hazard reductions with and without adjusting for the average treatment effect on UACR. The statistical supplement provides further details (Supplementary Note 3).

Application to future trials. From the trial-level meta-regression, we computed 95% Bayesian prediction intervals and estimated the probabilities of clinical benefit (defined as HR < 1.0 or HR < 0.8) for a large, moderate or small sized RCT corresponding to several possible treatment effects on UACR. By analogy with diagnostic testing, we refer to the estimated probability of clinical benefit at a particular value for the observed treatment effect on UACR as the trial-level positive predictive value (PPV_{trial}) corresponding to that observed treatment effect. For these calculations we assumed an s.d. in the change in log UACR of 0.75. We defined large, moderate and small RCTs as having 1,600 total (800 per group), 800 (400 per group) and 200 (100 per group) participants, respectively.

We also computed the threshold associated with the smallest observed treatment effect on UACR that would assure a high probability of benefit of the treatment on the clinical endpoint, defined as a $\text{PPV}_{\text{trial}}$ of 97.5%.

Prior distributions. We fit the second-stage model using Markov chain Monte Carlo (MCMC) sampling. For all model fitting, we used diffuse prior distributions for the model parameters. Across analyses, the priors for the mean treatment effects on log ACR, the meta-regression intercept and the meta-regression slope each were taken to be highly diffuse normal distributions with mean zero. For the primary analysis on the overall set of studies and for models where the intercept and slope varied by CKD severity variables, the variance of the prior was 10,000. To constrain the sampler in the presence of limited data within subgroups, the between-subgroup mean true treatment effects, intercept and slope terms had normal priors with variance of 25, which we still consider to be highly diffuse in this application. Throughout analyses using different meta-regression models, the priors for the variances of the treatment effects on the clinical endpoint (residual variance) and on change in log ACR were each taken to be inverse gamma distributions with shape parameter 0.261 and scale parameter 0.000408. This weakly informative prior distribution was selected by the investigators to assign 1/3 prior probabilities each to low treatment effect heterogeneity (which we defined as a treatment effect s.d. on the log scale ≤ 0.05), medium treatment effect heterogeneity (defined as a treatment effect s.d. on the log scale between 0.05 and 0.20) and high treatment effect heterogeneity (defined as a treatment effect s.d. on the log scale > 0.20). Lastly, in fitting the partial pooling models for subgroup analyses, a between-subgroup normal distribution was assumed for each subgroup-specific parameter (for example, the within-subgroup meta-regression slope). The variances of these half-normal priors were set to 25. The priors for other secondary analyses are described in the statistical supplement (Supplementary Note 3).

Analyses were performed using SAS version 9.4 (SAS Institute), R 3.16.1 (R Project for Statistical Computing; <https://www.r-project.org/>)³⁶ and RStan 2.21.5 (ref. 37)

Reporting summary

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

Data availability

All data used in the analysis were obtained by the CKD-EPI CT group through third parties. Data use agreements prohibit the CKD-EPI CT group from sharing data with parties external to the agreement. See Supplementary Note 1 for the identity of third-party providers. The following datasets can be requested through data-sharing platforms: Vivli: CANVAS ([NCT01032629](#)), CANVAS-R ([NCT01989754](#)), CREDENCE ([NCT02065791](#)), EMPA-REG Outcome ([NCT01131676](#)), FIDELIO-DKD ([NCT02540993](#)); NIDDK: AASK ([NCT04364139](#)), HALT-PKD A and B ([NCT00283686](#)), MDRD ([NCT03202914](#)); and the sponsor's website: LEADER ([NCT01179048](#)).

Code availability

The statistical code used for the primary analysis can be found at <https://github.com/UofUEpiBio/ckdepict>.

References

30. Kidney Disease: Improving Global Outcomes (KDIGO). KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. *Kidney Int. Suppl.* **3**, 1–150 (2013).
31. Levey, A. S. et al. A new equation to estimate glomerular filtration rate. *Ann. Intern. Med.* **150**, 604–612 (2009).
32. Collier, W., Haaland, B., Inker, L. & Greene, T. Handling missing within-study correlations in the evaluation of surrogate endpoints. *Stat. Med.* **42**, 4738–4762 (2023).
33. Prasad, V., Kim, C., Burotto, M. & Vandross, A. The strength of association between surrogate end points and survival in oncology: a systematic review of trial-level meta-analyses. *JAMA Intern. Med.* **175**, 1389–1398 (2015).
34. Collier, W., Haaland, B., Inker, L. A., Heerspink, H. J. L. & Greene, T. Comparing Bayesian hierarchical meta-regression methods and evaluating the influence of priors for evaluations of surrogate endpoints on heterogeneous collections of clinical trials. *BMC Med. Res. Methodol.* **24**, 39 (2024).
35. Collier, W. et al. Evaluation of variation in the performance of GFR slope as a surrogate end point for kidney failure in clinical trials that differ by severity of CKD. *Clin. J. Am. Soc. Nephrol.* **18**, 183–192 (2023).
36. Viechtbauer, W. Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* **36**, 1–48 (2010).
37. R Core Team *R: A Language and Environment for Statistical Computing* (R Core Team, 2022).

Acknowledgements

This work is presented on behalf of the CKD-EPI Clinical Trials (CKD-EPI CT) group. We thank all investigators, study teams and participants of the studies included in the meta-analysis. A full list of members and their affiliations appears in the Supplementary Information (Supplementary Note 1). We thank the National Kidney Foundation and its sponsors for funding this research. Support for CKD-EPI CT was obtained from multi-sponsor funding to the National Kidney Foundation. The National Kidney Foundation received consortium support from the following companies: Alexion, AstraZeneca, Bayer, Boehringer Ingelheim, CSL Behring, Novartis, Novo Nordisk A/S, ProKidney, Roche/Genentech and Travers Therapeutics. The funding sources had no role in the design and conduct of the study; the collection, management, analysis and interpretation of the data; the preparation, review or approval of the manuscript; the decision to submit the manuscript for publication; or any of the publications where the original results were reported. This work also received support from the Utah Study Design and Biostatistics Center, with funding in part from the National Center for Advancing Translational Sciences of the National Institutes of Health under award number UL1TR002538. The support and resources from the Center for High Performance Computing at the University of Utah are also gratefully acknowledged. The HALT-PKD Study A and Study B were conducted by the investigators of the respective studies and supported by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). The data from these studies reported here were supplied by the NIDDK Central Repository. This paper does not necessarily reflect the opinions or views of the studies, the NIDDK Central Repository or the NIDDK. We are grateful for data from AstraZeneca, Bayer AG, Boehringer Ingelheim and Novo Nordisk A/S.

Author contributions

H.J.L.H., T.G. and L.A.I. conceptualized the work and acquired funding for it. H.J.L.H., T.G., L.A.I., G.B.A., F.C.-F., J.F., T.H., E.I., T.H.J., J.B.L., P.K.T.L., F.L., B.D.M., B.L.N., V.P., R.D.P., F.P.S., C.W. and J.C. curated the data. T.G., W.H.C., H.T. and J.C. performed formal analyses. S.M. and W.H.C. visualized the results. H.J.L.H. wrote the original draft of the manuscript. All authors reviewed and edited the draft.

Competing interests

H.J.L.H. has served as a consultant for AstraZeneca, Alexion, Amgen, Bayer, Boehringer Ingelheim, Biocity Biopharmaceutics, Dimerix, Eli Lilly, Novartis, Novo Nordisk, Roche and Travers Therapeutics and has received grant support from AstraZeneca, Bayer, Boehringer

Ingelheim, Janssen and Novo Nordisk. W.H.C. reports funding from the National Kidney Foundation to his institute for his graduate research assistantship. G.B.A. reports funding for research studies through Columbia University, Calliditas, Travere Therapeutics, Novartis, Equillium, Roche, Sanofi, Reata, Apellis, Nephronet, Goldfinch and Vertex; consultancy for Sanofi-Genzyme, Roche/Genentech, Apellis, Novartis, Travere Therapeutics, Calliditas, GlaxoSmithKline and Aurinia; and speaker for GlaxoSmithKline and Aurinia, both on lupus nephritis. J.F. reports lecture and consulting honoraria from Alnylam, Amgen, AstraZeneca, Bayer, Boehringer Ingelheim, Calliditas, Fresenius, Novartis, Omeros, Travere Therapeutics and Vifor and membership on the data safety monitoring board in trials for Novo Nordisk A/S and Visterra. E.I. received research funding and honoraria for lectures from AstraZeneca, Tanabe Mitsubishi, Boehringer Ingelheim, Daiichi-Sankyo, Kyowa Kirin and Novo Nordisk A/S. J.B.L. is a consultant for CSL Behring and BioVie. P.K.T.L. received speaker honoraria from FibroGen, AstraZeneca and Baxter. B.D.M. is the National Leader in the ASCEND-D and ASCEND-ND trials of GlaxoSmithKline and the PROTECT and DUPLEX trials of Travere Therapeutics as well as a steering committee member in the Novartis APPLAUSE trial and the Vera Therapeutics ATACICEPT trial. B.L.N. has received fees for travel support, advisory boards, scientific presentations and steering committee roles from AstraZeneca, Alexion, Bayer, Boehringer Ingelheim, Cambridge Healthcare Research, Cornerstone Medical Education, the Limbic, Janssen, Medscape, Novo Nordisk and Travere Therapeutics, with all honoraria paid to The George Institute for Global Health. R.D.P. has been on steering committees for AbbVie* and Vertex*; has consulted for Otsuka*, Retex*, Estuary, SouffleTx, Sanofi-Genzyme, Regenephro and Regulus*; and is the section editor for cystic kidney disease for UpToDate (*paid to Tufts Medical Center). F.P.S. reports funding from

the Italian Ministry of Health and the Italian Ministry of University and Research. C.W. reports lecture and consulting honoraria from Alexion, AstraZeneca, Bayer, Boehringer Ingelheim, CSL Vifor, GlaxoSmithKline, Novartis, Novo Nordisk and VeraTX. T.G. reports funding from the National Institutes of Health, the Patient Center Outcome Research Institute, the National Kidney Foundation and the Collaborative Study Group. L.A.I. reports funding to Tufts Medical Center from the National Institutes of Health, the National Kidney Foundation, Chinook, R3R, Otsuka, Alexion and Novartis; consulting agreements to Tufts Medical Center from Alexion and AstraZeneca; and consulting agreements from Dimerix. The other authors declare no competing interests.

Additional information

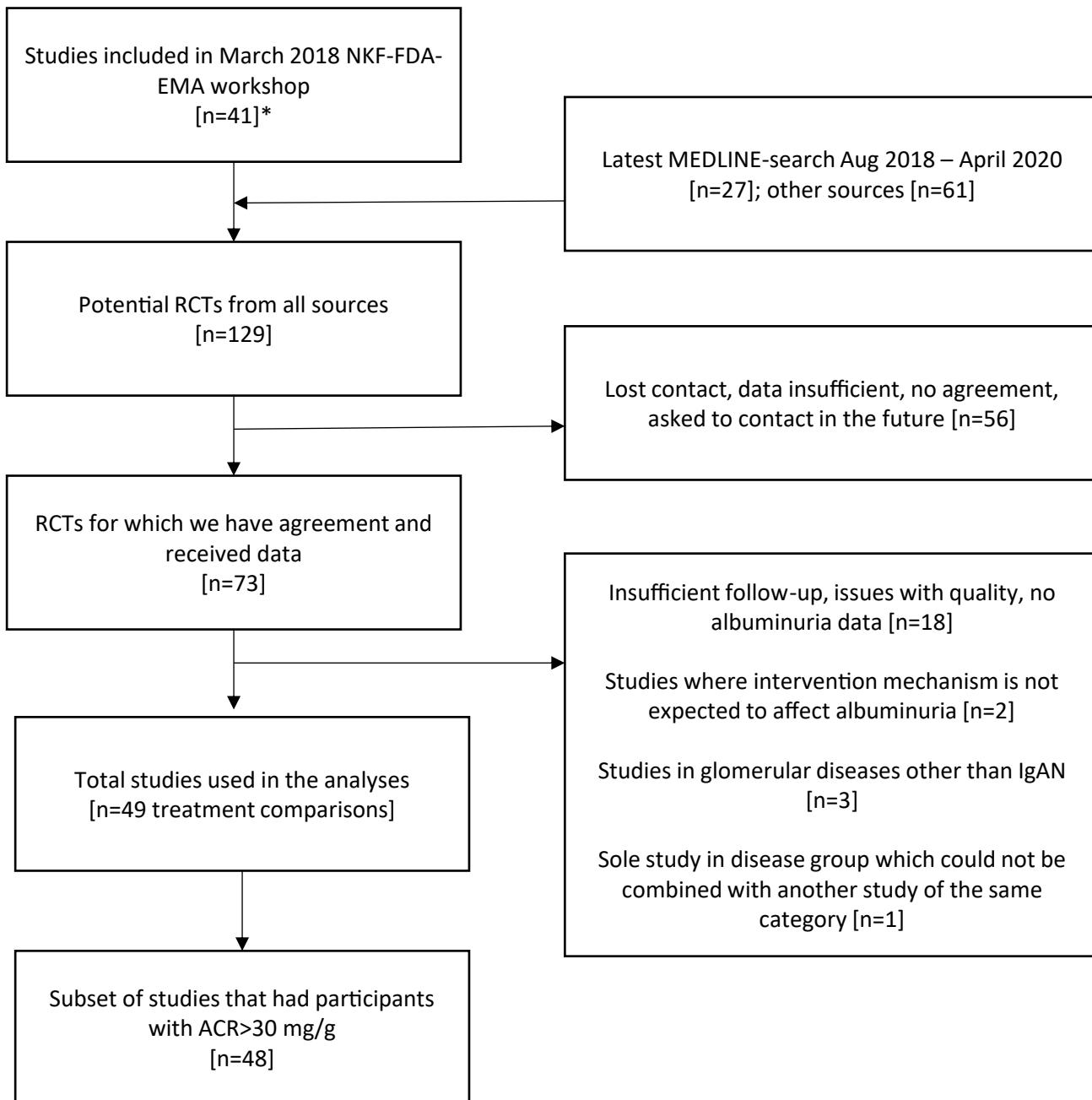
Extended data is available for this paper at
<https://doi.org/10.1038/s41591-025-04057-z>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41591-025-04057-z>.

Correspondence and requests for materials should be addressed to Hiddo J. L. Heerspink.

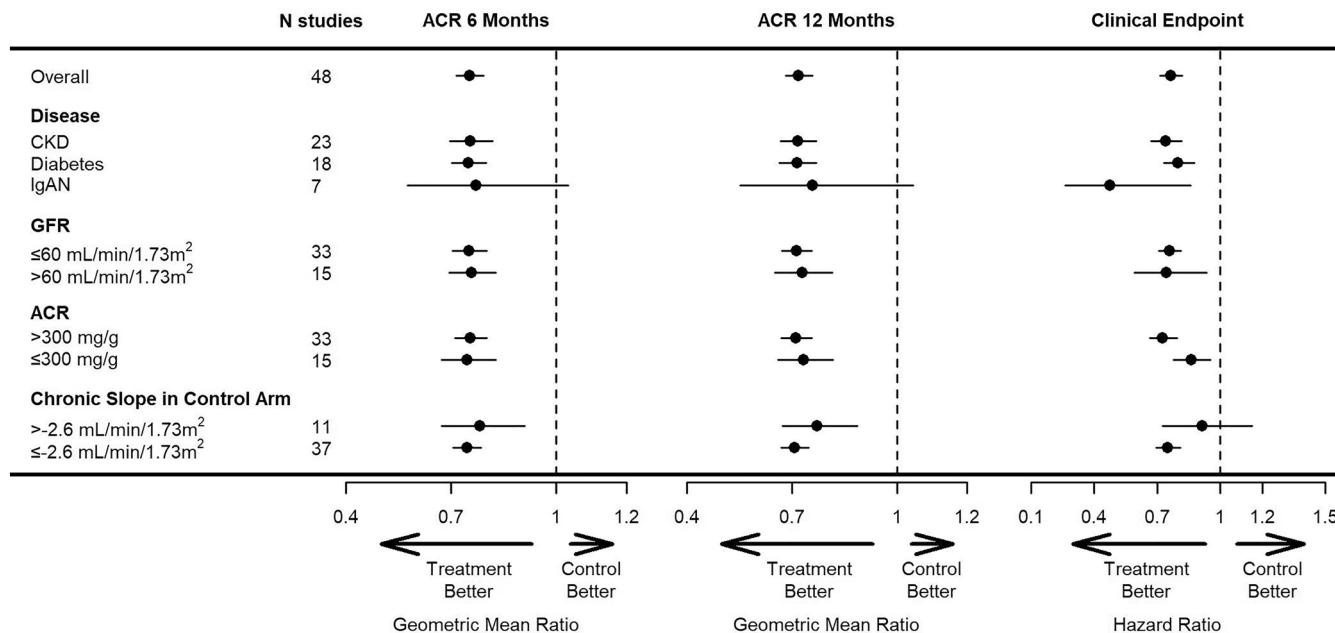
Peer review information *Nature Medicine* thanks Petter Bjornstad, Jennifer Lees, Johannes Mann and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Michael Basson, in collaboration with the *Nature Medicine* team.

Reprints and permissions information is available at
www.nature.com/reprints.



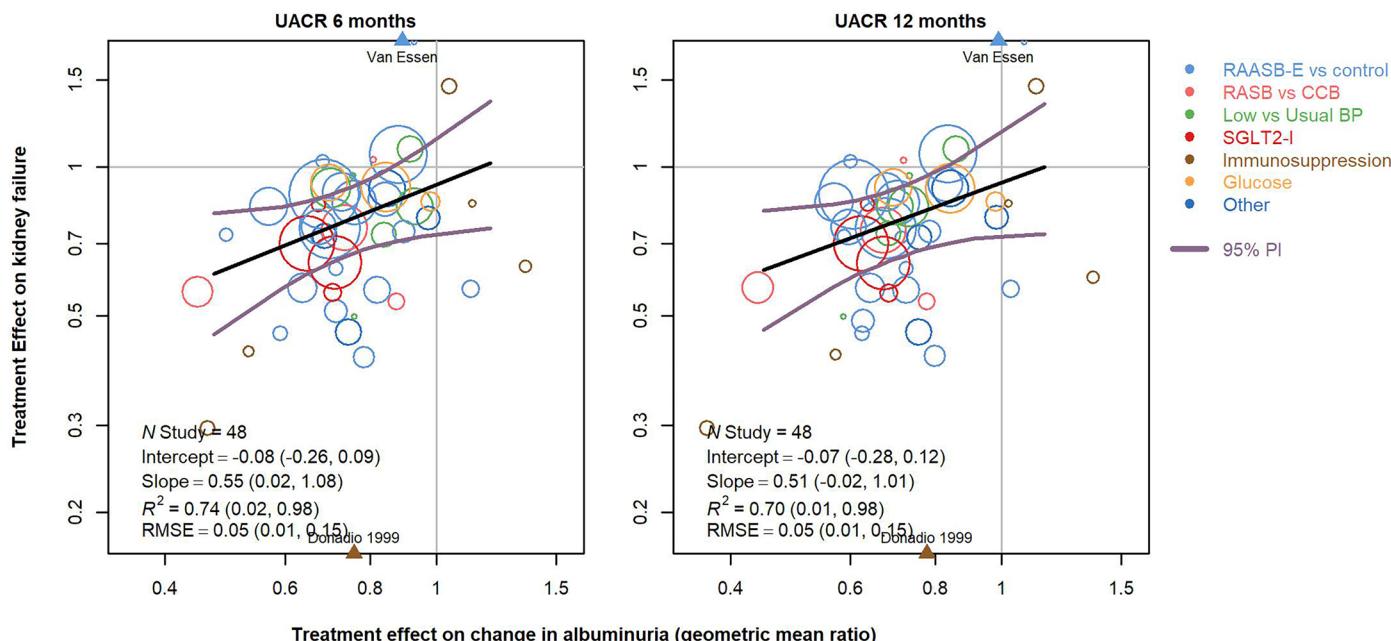
Extended Data Fig. 1 | Flowchart for study inclusion. We included studies only if the first follow-up serum creatinine was at 12 months or earlier, and if there was follow-up data for more than 12 months after this measurement. We require a specified number of clinical events for inclusion. [For glomerular disease, this threshold was 10 events, for kidney disease it was a follow-up 500 or more person-years and 30 or more events, for high-risk populations (diabetes, hypertension, cardiovascular disease, heart failure - not selected for having kidney disease), it was a follow-up 1000 or more person-years and 30 or more events]. *This was

published earlier¹⁰, and studies were divided by treatment comparison (8 studies had multiple arms) and pooling small studies <100 participants if the disease and intervention were the same (9 studies pooled into 4). For a sensitivity analysis, two more studies^{17,18} were included in the final database using aggregate results shared with us by representatives from those studies. UACR, urine albumin:creatinine ratio; IgAN, IgA nephropathy; NKF-FDA-EMA; National Kidney Foundation – US Food and Drug Administration – European Medicines Agency; RCT, randomized controlled trial.

**Extended Data Fig. 2 | Treatment effect on UACR and on the clinical endpoint.**

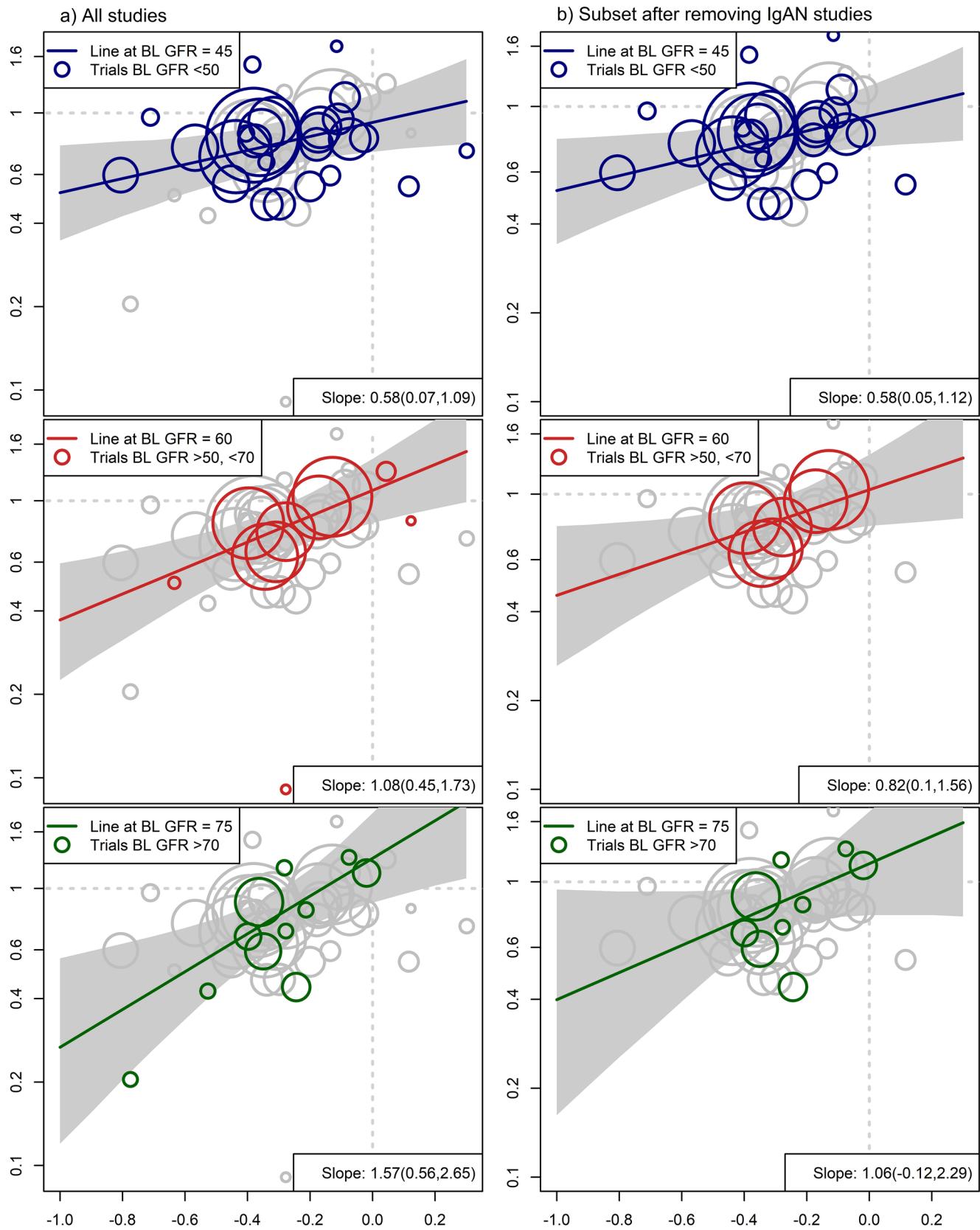
Treatment effects on UACR are expressed as geometric mean ratios and were estimated by performing analyses of covariance within each study. Treatment effects on clinical endpoints are expressed as hazard ratio and were estimated by performing Cox proportional hazard regression within each study. The figure presents estimated pooled average treatment effects based on random effects

meta-analysis, with 95% confidence intervals, for groups of studies classified by their type of disease, mean baseline GFR, median baseline UACR, and mean chronic slope in the control arm. UACR, urine albumin:creatinine ratio; CKD, chronic kidney disease; GFR, glomerular filtration rate; IgAN, IgA nephropathy; N, number of.



Extended Data Fig. 3 | Trial level association between change in UACR and kidney failure alone. The treatment effects on change in albuminuria are expressed on the horizontal axis. The treatment effects on the clinical endpoint (that is, kidney failure defined as GFR < 15 mL/min/1.73m², initiation of chronic treatment with dialysis, or kidney transplantation) are expressed on the vertical axis. Treatment effects on UACR are expressed as geometric mean ratios. Each circle is a separate study with the size of the circle proportional to the number of events. The colors of the circles indicate intervention type. The black line is

the line of the regression through the studies. The plum-colored line represents the Bayesian 95% prediction band. BP, Blood Pressure; CCB, Calcium Channel Blocker; HR, hazard ratio; N, number of; R², coefficient of determination; RAASB-E, renin-angiotensin-aldosterone system blockade or endothelial receptor antagonist; RASB, renin-angiotensin system blocker; RMSE, root mean squared error; SGLT2i, Sodium-Glucose Cotransporter-2 inhibitor; UACR, urine albumin:creatinine ratio.

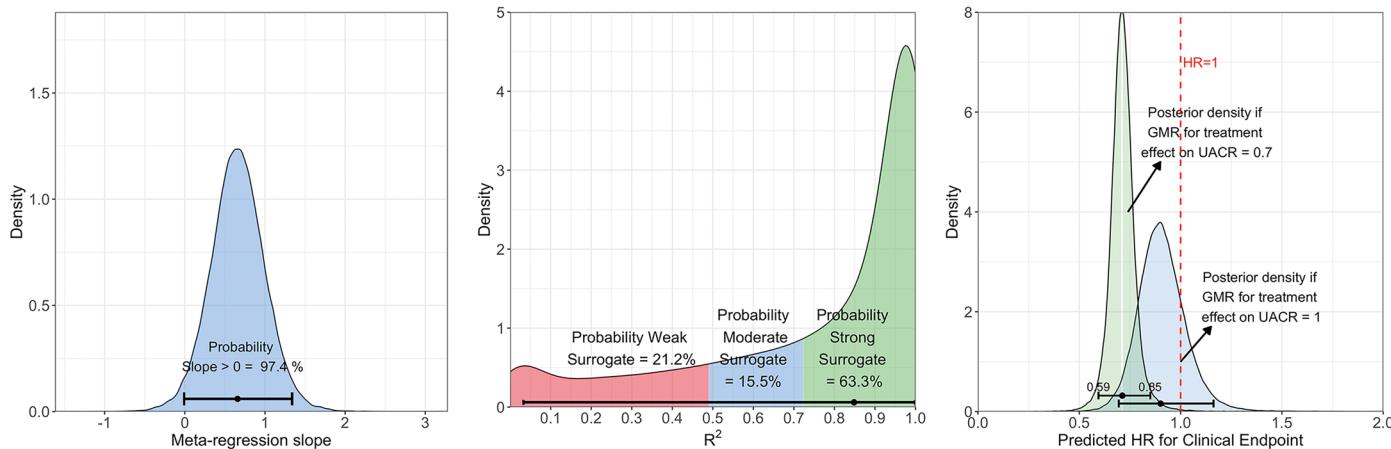


Extended Data Fig. 4 | See next page for caption.

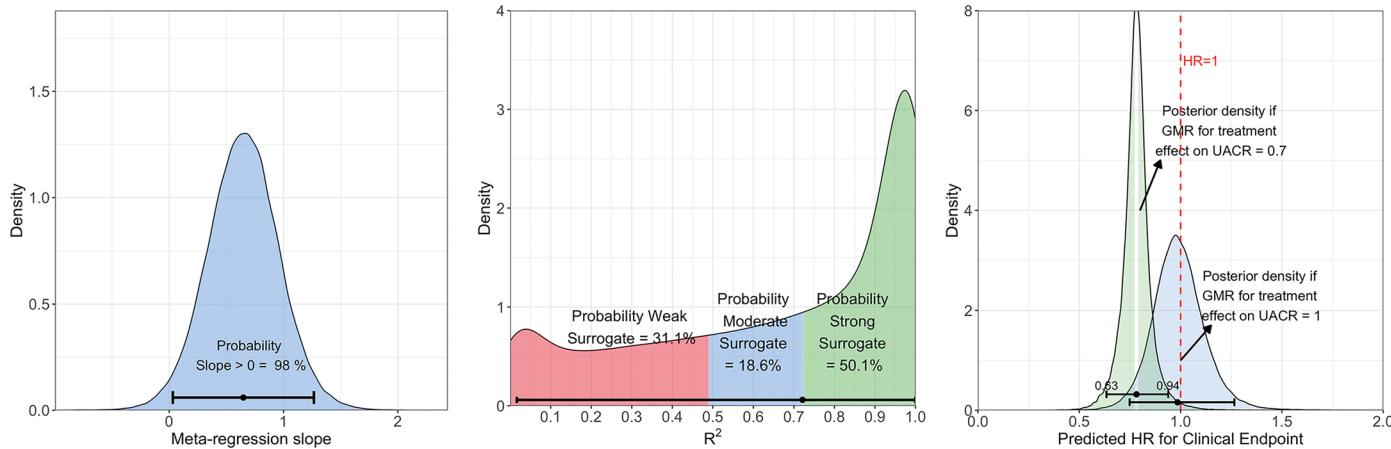
Extended Data Fig. 4 | Impact of baseline GFR on the trial-level association between UACR and the clinical endpoint. **a**) all studies, **b**) subset after removing IgAN studies. These figures display the meta-regression line relating treatment effects on the clinical endpoint to treatment effects on 6-month change in UACR from the model where the meta-regression intercept and slope are allowed to vary continuously as a function of mean baseline GFR. The meta-regression line is displayed as a function of chosen candidate values of mean baseline GFR, and the highlighted trials are those within a range of the value chosen to plot the line in

each subplot. The numerical summary displayed within each subplot (bottom-right) is the posterior median and, within parentheses, the posterior 2.5th and 97.5th percentile for the meta-regression slope at the specified baseline GFR level. The regression line uses the posterior median for relevant meta-regression parameters (for example, intercept and slopes) and the curved, shaded bands represent the 95% prediction bands for the true treatment effect on the clinical endpoint in an individual trial. BL, baseline.

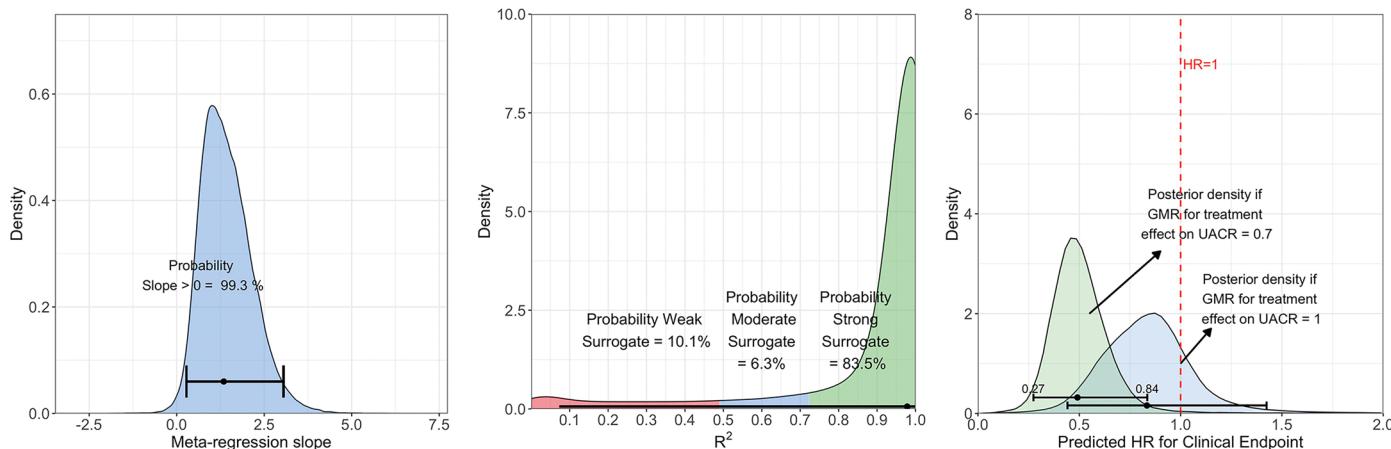
A) Chronic kidney disease



B) Diabetes



C) Immunoglobulin A nephropathy



Extended Data Fig. 5 | Posterior distributions for three aspects of the trial-level association between the treatment effects on urinary albumin-to-creatinine ratio (UACR) at 6 months and the clinical endpoint by disease subgroups. a) Chronic kidney disease, b) Diabetes; c) immunoglobulin A nephropathy. The graphs show the full posterior distributions of the meta-

regression slope and R^2 (left and middle panels), and the posterior predictive distribution of true treatment effects on the clinical endpoint given selected true treatment effects on UACR (right panel). The black intervals in the lower part of each panel extend from the 2.5th to the 97.5th percentile of the relevant posterior distribution.

Extended Data Table 1 | Trial-level analysis for change in UACR and kidney failure alone

| Group | Severity Level | Meta-Regression Slope | Intercept | R ² | RMSE |
|--|--------------------|--|---|---|---|
| Overall | | 0.55 (0.02, 1.08) | -0.08 (-0.26, 0.09) | 0.74 (0.02, 0.98) | 0.05 (0.01, 0.15) |
| Baseline GFR (ml/min/1.73m ²) ¹ | 45 75 | 0.55(-0.01,1.07) 1.31(0.13,2.54) | -0.08(-0.27,0.08) 0.19(-0.19,0.57) | 0.77(0.02,0.99) 0.95(0.17,1.00) | 0.04(0.01,0.14) 0.04(0.01,0.14) |
| Baseline UACR (mg/g) ¹ | 100 300 1000 | 0.67(-0.38,1.76) 0.57(-0.01,1.16) 0.44(-0.30,1.19) | 0.04(-0.31,0.39) -0.05(-0.24,0.14) -0.14(-0.39,0.10) | 0.82(0.01,0.99) 0.76(0.02,0.98) 0.65(0.00,0.98) | 0.05(0.01,0.15) 0.05(0.01,0.15) 0.05(0.01,0.15) |
| GFR Rate of Progression (ml/min/1.73m ² /yr) ^{1,2} | -5 -3 -1 | 0.71(0.03,1.42) 0.41(-0.30,1.12) 0.11(-1.41,1.62) | -0.03(-0.25,0.19) -0.13(-0.36,0.09) -0.23(-0.69,0.22) | 0.81(0.03,0.99) 0.59(0.00,0.98) 0.69(0.00,0.99) | 0.05(0.02,0.16) 0.05(0.02,0.16) 0.05(0.02,0.16) |

Extended Data Table 2 | Analysis of 6-month change in UACR with continuous interaction by disease severity

| Parameter | Interaction Variable: Mean Baseline GFR (per 10 mL/min per 1.73m ² higher) | Interaction Variable: Median Baseline UACR (per 1 log-transformed mg/g) | Interaction Variable: Control Arm Chronic Slope (per 1 mL/min per 1.73m ² /year higher) |
|---|--|--|---|
| Meta-Regression intercept at mean level for each severity measure | -0.02 (-0.19, 0.14) | -0.06 (-0.23, 0.11) | -0.05 (-0.23, 0.13) |
| Meta-Regression slope at mean level for each severity measure | 0.75 (0.26, 1.26) | 0.61 (0.10, 1.14) | 0.70 (0.15, 1.24) |
| Change in Meta-Regression Intercept for a Designated Change in Severity (Interaction Intercept) | 0.11 (0.01, 0.22) | -0.12 (-0.31, 0.07) | -0.02 (-0.16, 0.11) |
| Change in Meta-Regression Slope for a Designated Change in Severity (Interaction Slope) | 0.33 (0.00, 0.69) | -0.18 (-0.78, 0.41) | -0.06 (-0.55, 0.41) |
| RMSE | 0.06 (0.02, 0.17) | 0.07 (0.02, 0.17) | 0.09 (0.02, 0.19) |

Extended Data Table 3 | Application of change in UACR as surrogate endpoint in new randomized clinical trials: predicted treatment effect on clinical endpoint and $\text{PPV}_{\text{trial}}$ for change in albuminuria at 12 months

| Observed Treatment effect on change in UACR | RCT of 1600 | | RCT of 800 | | RCT of 200 | |
|---|--|-----------------------------|--|-----------------------------|--|-----------------------------|
| | Median HR and 95% Bayesian Prediction Interval | $\text{PPV}_{\text{trial}}$ | Median HR and 95% Bayesian Prediction Interval | $\text{PPV}_{\text{trial}}$ | Median HR and 95% Bayesian Prediction Interval | $\text{PPV}_{\text{trial}}$ |
| | 0.5 | 0.63 (0.46, 0.82) | 1.00 | 0.63 (0.46, 0.82) | 1.00 | 0.63 (0.45, 0.83) |
| 0.6 | 0.71 (0.54, 0.88) | 1.00 | 0.71 (0.54, 0.88) | 1.00 | 0.71 (0.52, 0.90) | 0.99 |
| 0.7 | 0.77 (0.61, 0.96) | 0.99 | 0.77 (0.60, 0.96) | 0.99 | 0.77 (0.59, 0.99) | 0.98 |
| 0.8 | 0.84 (0.65, 1.05) | 0.95 | 0.84 (0.65, 1.06) | 0.94 | 0.83 (0.64, 1.09) | 0.92 |
| 0.9 | 0.90 (0.69, 1.16) | 0.83 | 0.90 (0.69, 1.16) | 0.82 | 0.89 (0.68, 1.20) | 0.80 |
| 1 | 0.96 (0.72, 1.27) | 0.63 | 0.96 (0.72, 1.28) | 0.63 | 0.95 (0.71, 1.32) | 0.64 |
| Threshold to assure $\text{PPV}_{\text{trial}} \geq 97.5\%$ | | 0.73 | | 0.73 | | 0.71 |

Extended Data Table 4 | Application of change in UACR as surrogate endpoint in new randomized controlled trials: PPV_{trial} for change in UACR

| Observed Treatment effect on change in UACR | Change in UACR at 6 months | | | | | | Change in UACR at 12 months | | | | | |
|---|----------------------------|----------|------------|----------|------------|----------|-----------------------------|----------|------------|----------|------------|----------|
| | RCT of 1600 | | RCT of 800 | | RCT of 200 | | RCT of 1600 | | RCT of 800 | | RCT of 200 | |
| | HR < 1 | HR < 0.8 | HR < 1 | HR < 0.8 | HR < 1 | HR < 0.8 | HR < 1 | HR < 0.8 | HR < 1 | HR < 0.8 | HR < 1 | HR < 0.8 |
| Clinical endpoint | | | | | | | | | | | | |
| 0.5 | 1.00 | 0.98 | 1.00 | 0.98 | 1.00 | 0.98 | 1.00 | 0.97 | 1.00 | 0.97 | 1.00 | 0.96 |
| 0.6 | 1.00 | 0.95 | 1.00 | 0.95 | 1.00 | 0.92 | 1.00 | 0.90 | 1.00 | 0.89 | 0.99 | 0.87 |
| 0.7 | 0.99 | 0.79 | 0.99 | 0.78 | 0.99 | 0.73 | 0.99 | 0.65 | 0.99 | 0.65 | 0.98 | 0.63 |
| 0.8 | 0.97 | 0.38 | 0.97 | 0.40 | 0.95 | 0.43 | 0.95 | 0.31 | 0.94 | 0.32 | 0.92 | 0.35 |
| 0.9 | 0.88 | 0.16 | 0.87 | 0.17 | 0.84 | 0.21 | 0.83 | 0.15 | 0.82 | 0.16 | 0.80 | 0.19 |
| 1 | 0.67 | 0.08 | 0.66 | 0.08 | 0.66 | 0.10 | 0.63 | 0.09 | 0.63 | 0.09 | 0.64 | 0.11 |
| Kidney failure alone | | | | | | | | | | | | |
| 0.5 | 1.00 | 0.97 | 1.00 | 0.97 | 1.00 | 0.97 | 1.00 | 0.96 | 1.00 | 0.96 | 1.00 | 0.95 |
| 0.6 | 1.00 | 0.94 | 1.00 | 0.94 | 1.00 | 0.92 | 1.00 | 0.91 | 1.00 | 0.91 | 1.00 | 0.87 |
| 0.7 | 1.00 | 0.80 | 1.00 | 0.78 | 0.99 | 0.73 | 1.00 | 0.68 | 1.00 | 0.66 | 0.99 | 0.64 |
| 0.8 | 0.99 | 0.39 | 0.99 | 0.40 | 0.97 | 0.44 | 0.98 | 0.30 | 0.98 | 0.32 | 0.96 | 0.36 |
| 0.9 | 0.94 | 0.17 | 0.93 | 0.17 | 0.90 | 0.22 | 0.90 | 0.16 | 0.90 | 0.16 | 0.87 | 0.19 |
| 1 | 0.79 | 0.10 | 0.78 | 0.10 | 0.76 | 0.12 | 0.74 | 0.10 | 0.74 | 0.10 | 0.73 | 0.12 |

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 Data were managed in SAS version 9.4

Data analysis The statistical code we used for the primary analysis can be found at <https://github.com/UofUEpiBio/ckdepict>. Analyses were performed using SAS version 9.4 (SAS Institute, Cary, NC) and R 3.16.1 (R Project for Statistical Computing www.r-project.org) and RStan 2.21.5

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](#)

All data used in the analysis were obtained by the CKD-EPI CT group through third parties.

Data use agreements prohibit CKD-EPI CT from sharing data with parties external to the agreement.

See Appendix 1 for identity of third-party providers. The following datasets could be requested through data sharing platforms:
 Vivi: CANVAS (NCT01032629), CANVAS-R (NCT01989754), CREDENCE (NCT02065791), EMPA-REG Outcome (NCT01131676),
 FIDELIO-DKD (NCT02540993); NIDDK: AASK (NCT04364139), HALT-PKD A and B (NCT00283686), MDRD (NCT03202914);
 Sponsors' website: LEADER (NCT01179048)

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 We have used the term sex for biological sex of participants. We have also presented results of our main analysis by sex subgroups.

Reporting on race, ethnicity, or other socially relevant groupings

Not applicable

Population characteristics

We have provided descriptive tables for study participants in Table 1. This includes age, sex, diabetes status, and kidney function.

Recruitment

Participants were recruited in individual trials in the meta-analysis as per the protocol of that trial.
We have provided a risk of bias summary for selection of each trial in the meta-analysis.

Ethics oversight

Individual studies were reviewed by the participating centers' institutional review boards.
Tufts Medical Center Institutional Review Board deemed these analyses exempt from review.

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

Life sciences study design

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

Sample size We pre-specified the minimum number of participants, the amount of events, and years of follow-up needed in each type of study included in the meta-analysis. This information is available in Extended Data Figure 1.

Data exclusions We excluded studies as per criteria in Extended Data Figure 1

Replication We provide detailed methods for each aspect of the meta-analysis including the systematic review of studies to the statistical methods for obtaining the main statistics and uncertainty around the statistics. Results were reproducible with reasonable precision by changing number of iterations of Monte Carlo sampling

Randomization Randomization was performed by the original investigators of the studies. We assessed randomization of participants within individual studies included in the meta-analysis. This information is available in Supplementary Figure 1.

Blinding Blinding was not directly relevant as this a meta-analysis. But we assessed blinding of participants and investigators for individual \ studies included in the meta-analysis. This information is available in Supplementary Figure 1.

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 type="checkbox"/> | Antibodies |
| <input type="checkbox"/> | Eukaryotic cell lines |
| <input type="checkbox"/> | Palaeontology and archaeology |
| <input type="checkbox"/> | Animals and other organisms |
| <input type="checkbox"/> | Clinical data |
| <input type="checkbox"/> | Dual use research of concern |
| <input type="checkbox"/> | Plants |

Methods

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

Antibodies

Antibodies used

Describe all antibodies used in the study; as applicable, provide supplier name, catalog number, clone name, and lot number.

Validation

Describe the validation of each primary antibody for the species and application, noting any validation statements on the manufacturer's website, relevant citations, antibody profiles in online databases, or data provided in the manuscript.

Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

Cell line source(s)

State the source of each cell line used and the sex of all primary cell lines and cells derived from human participants or vertebrate models.

Authentication

Describe the authentication procedures for each cell line used OR declare that none of the cell lines used were authenticated.

Mycoplasma contamination

Confirm that all cell lines tested negative for mycoplasma contamination OR describe the results of the testing for mycoplasma contamination OR declare that the cell lines were not tested for mycoplasma contamination.

Commonly misidentified lines
(See [ICLAC register](#))

Name any commonly misidentified cell lines used in the study and provide a rationale for their use.

Palaeontology and Archaeology

Specimen provenance

Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information). Permits should encompass collection and, where applicable, export.

Specimen deposition

Indicate where the specimens have been deposited to permit free access by other researchers.

Dating methods

If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.

Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Ethics oversight

Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.

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

Animals and other research organisms

Policy information about [studies involving animals; ARRIVE guidelines](#) recommended for reporting animal research, and [Sex and Gender in Research](#)

Laboratory animals

For laboratory animals, report species, strain and age OR state that the study did not involve laboratory animals.

Wild animals

Provide details on animals observed in or captured in the field; report species and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.

Reporting on sex

Indicate if findings apply to only one sex; describe whether sex was considered in study design, methods used for assigning sex.

Provide data disaggregated for sex where this information has been collected in the source data as appropriate; provide overall

numbers in this Reporting Summary. Please state if this information has not been collected. Report sex-based analyses where performed, justify reasons for lack of sex-based analysis.

Field-collected samples

For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.

Ethics oversight

Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.

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

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

Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.

Study protocol

Note where the full trial protocol can be accessed OR if not available, explain why.

Data collection

Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.

Outcomes

Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.

Dual use research of concern

Policy information about [dual use research of concern](#)

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

- | | |
|--------------------------|---|
| No | Yes |
| <input type="checkbox"/> | <input type="checkbox"/> Public health |
| <input type="checkbox"/> | <input type="checkbox"/> National security |
| <input type="checkbox"/> | <input type="checkbox"/> Crops and/or livestock |
| <input type="checkbox"/> | <input type="checkbox"/> Ecosystems |
| <input type="checkbox"/> | <input type="checkbox"/> Any other significant area |

Experiments of concern

Does the work involve any of these experiments of concern:

No

Yes

- | | |
|--------------------------|--|
| <input type="checkbox"/> | <input type="checkbox"/> Demonstrate how to render a vaccine ineffective |
| <input type="checkbox"/> | <input type="checkbox"/> Confer resistance to therapeutically useful antibiotics or antiviral agents |
| <input type="checkbox"/> | <input type="checkbox"/> Enhance the virulence of a pathogen or render a nonpathogen virulent |
| <input type="checkbox"/> | <input type="checkbox"/> Increase transmissibility of a pathogen |
| <input type="checkbox"/> | <input type="checkbox"/> Alter the host range of a pathogen |
| <input type="checkbox"/> | <input type="checkbox"/> Enable evasion of diagnostic/detection modalities |
| <input type="checkbox"/> | <input type="checkbox"/> Enable the weaponization of a biological agent or toxin |
| <input type="checkbox"/> | <input type="checkbox"/> Any other potentially harmful combination of experiments and agents |

Plants

| | |
|-----------------------|---|
| Seed stocks | Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures. |
| Novel plant genotypes | Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied. |
| Authentication | Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosaicism, off-target gene editing) were examined. |

ChIP-seq

Data deposition

- Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).
- Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links

May remain private before publication.

For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.

Files in database submission

Provide a list of all files available in the database submission.

Genome browser session (e.g. [UCSC](#))

Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.

Methodology

Replicates

Describe the experimental replicates, specifying number, type and replicate agreement.

Sequencing depth

Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.

Antibodies

Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.

Peak calling parameters

Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.

Data quality

Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.

Software

Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.

Flow Cytometry

Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.

Instrument

Identify the instrument used for data collection, specifying make and model number.

Software

Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.

Cell population abundance

Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.

Gating strategy

Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type

Indicate task or resting state; event-related or block design.

Design specifications

Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.

Behavioral performance measures

State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).

Acquisition

Imaging type(s)

Specify: functional, structural, diffusion, perfusion.

Field strength

Specify in Tesla

Sequence & imaging parameters

Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.

Area of acquisition

State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.

Diffusion MRI

Used

Not used

Preprocessing

Preprocessing software

Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).

Normalization

If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.

Normalization template

Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.

Noise and artifact removal

Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).

Volume censoring

Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.

Statistical modeling & inference

Model type and settings

Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).

Effect(s) tested

Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.

Specify type of analysis: Whole brain ROI-based Both

Statistic type for inference

Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.

(See [Eklund et al. 2016](#))

Correction

Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).

Models & analysis

n/a Involved in the study

Functional and/or effective connectivity

Graph analysis

Multivariate modeling or predictive analysis

Functional and/or effective connectivity

Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).

Graph analysis

Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).

Multivariate modeling and predictive analysis

Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.