Change in albuminuria as a surrogate endpoint for progression of kidney disease: a meta-analysis of treatment effects in randomised clinical trials

*Hiddo J L Heerspink, Tom Greene, Hocine Tighiouart, Ron T Gansevoort, Josef Coresh, Andrew L Simon, Tak Mao Chan, Fan Fan Hou, Julia B Lewis, Francesco Locatelli, Manuel Praga, Francesco Paolo Schena, Andrew S Levey, Lesley A Inker, for the Chronic Kidney Disease Epidemiology Collaboration**

Summary

Background Change in albuminuria has strong biological plausibility as a surrogate endpoint for progression of chronic kidney disease, but empirical evidence to support its validity is lacking. We aimed to determine the association between treatment effects on early changes in albuminuria and treatment effects on clinical endpoints and surrogate endpoints, to inform the use of albuminuria as a surrogate endpoint in future randomised controlled trials.

Methods In this meta-analysis, we searched PubMed for publications in English from Jan 1, 1946, to Dec 15, 2016, using search terms including "chronic kidney disease", "chronic renal insufficiency", "albuminuria", "proteinuria", and "randomized controlled trial"; key inclusion criteria were quantifiable measurements of albuminuria or proteinuria at baseline and within 12 months of follow-up and information on the incidence of end-stage kidney disease. We requested use of individual patient data from the authors of eligible studies. For all studies that the authors agreed to participate and that had sufficient data, we estimated treatment effects on 6-month change in albuminuria and the composite clinical endpoint of treated end-stage kidney disease, estimated glomerular filtration rate of less than 15 mL/min per 1·73 m², or doubling of serum creatinine. We used a Bayesian mixed-effects meta-regression analysis to relate the treatment effects on albuminuria to those on the clinical endpoint across studies and developed a prediction model for the treatment effect on the clinical endpoint on the basis of the treatment effect on albuminuria.

Findings We identified 41 eligible treatment comparisons from randomised trials (referred to as studies) that provided sufficient patient-level data on 29 979 participants (21 206 [71%] with diabetes). Over a median follow-up of 3·4 years (IQR 2·3–4·2), 3935 (13%) participants reached the composite clinical endpoint. Across all studies, with a meta-regression slope of 0·89 (95% Bayesian credible interval [BCI] 0·13–1·70), each 30% decrease in geometric mean albuminuria by the treatment relative to the control was associated with an average 27% lower hazard for the clinical endpoint (95% BCI 5–45%; median R^2 0·47, 95% BCI 0·02–0·96). The association strengthened after restricting analyses to patients with baseline albuminuria of more than 30 mg/g (ie, 3·4 mg/mmol; R^2 0·72, 0·05–0·99]). For future trials, the model predicts that treatments that decrease the geometric mean albuminuria to 0·7 (ie, 30% decrease in albuminuria) relative to the control will provide an average hazard ratio (HR) for the clinical endpoint of 0·68, and 95% of sufficiently large studies would have HRs between 0·47 and 0·95.

Interpretation Our results support a role for change in albuminuria as a surrogate endpoint for the progression of chronic kidney disease, particularly in patients with high baseline albuminuria; for patients with low baseline levels of albuminuria this association is less certain.

Funding US National Kidney Foundation.

Introduction

Chronic kidney disease is a major global health concern with few proven effective therapies. One of the challenges in the development and assessment of therapies for chronic kidney disease is that randomised controlled trials to assess the efficacy and safety of novel therapies traditionally use the composite clinical endpoint of end-stage kidney disease, estimated glomerular filtration rate (eGFR) of less than 15 mL/min per 1·73 m², and doubling of serum creatinine concentrations (equivalent to a 57% decrease in eGFR).[1] These endpoints are late events in the progression of chronic kidney disease. As a result, to obtain a sufficient number of outcome events, many randomised controlled trials have been restricted to patients at late stages of chronic kidney disease or with a rapid decrease in their glomerular filtration rate (GFR), thereby restricting the study populations and the feasibility of the trials. Moreover, appropriate intervention at earlier stages of chronic kidney disease has been advocated because early treatment might be more beneficial than interventions applied at later stages.[2] Therefore, alternative endpoints are needed to enable

Lancet Diabetes Endocrinol 2019

Published Online
January 8, 2019
http://dx.doi.org/10.1016/S2213-8587(18)30314-0

See Online/Comment
http://dx.doi.org/10.1016/S2213-8587(18)30352-8

*Members listed at the end of the Article

Department of Clinical Pharmacy and Pharmacology (H J L Heerspink PharmD) **and Department of Nephrology** (R T Gansevoort MD), **University of Groningen, University Medical Center Groningen, Groningen, Netherlands; Division of Epidemiology, Department of Internal Medicine, University of Utah, Salt Lake City, UT, USA** (T Greene PhD); **Institute for Clinical Research and Health Policy Studies** (H Tighiouart MS) **and Division of Nephrology** (A L Simon MS, Prof A S Levey MD, L A Inker MD), **Tufts Medical Center, Boston, MA, USA** (H Tighiouart); **Tufts Clinical and Translational Science Institute, Tufts University, Boston, MA, USA** (H Tighiouart); **Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA** (Prof J Coresh MD); **Department of Medicine, University of Hong Kong, Queen Mary Hospital, Pokfulam, Hong Kong** (T M Chan MD); **Division of Nephrology, Nanfang Hospital, Southern Medical University, Guangzhou, China** (F F Hou MD); **Division of Nephrology, Vanderbilt University, Nashville, TN, USA** (J B Lewis MD); **Department of Nephrology and Dialysis, Hospital Alessandro Manzoni, Lecco, Italy** (F Locatelli MD); **Nephrology Department, Hospital Universitario 12 de Octubre, Madrid, Spain**(M Praga MD); **and Renal, Dialysis and Transplant Unit,**

Department of Emergency and Organ Transplantation, University of Bari, Bari, Italy (F P Schena MD)

Correspondence to: Dr Lesley A Inker, Division of Nephrology, Tufts Medical Center, Boston, MA 02111, USA
linker@tuftsmedicalcenter.org

Research in context

Evidence before this study

The current endpoints for clinical trials of progression of chronic kidney disease are end-stage kidney disease and doubling of serum creatinine, which approximates a 57% decrease in estimated glomerular filtration rate. Interest is increasing in the use of alternative endpoints in clinical trials to shorten trial duration and decrease required sample sizes. Emerging data support a strong biological possibility for change in albuminuria as a surrogate endpoint for progression of chronic kidney disease. We searched PubMed for publications in English between Jan 1, 1946, and Dec 15, 2016, included the search terms "chronic kidney disease", "chronic renal insufficiency", "albuminuria", "proteinuria", and "randomized controlled trial". In observational studies, albuminuria has consistently been shown to be a strong predictor of end-stage kidney disease in a range of patients and settings. However, two previous meta-analyses reported conflicting results with respect to the strength of the association between treatment effects on early change in albuminuria and treatment effects on end-stage kidney disease. The US National Kidney Foundation, in collaboration with the US Food and Drug Administration and European Medicines Agency, sponsored a workshop to investigate candidate surrogate endpoints for clinical trials of drugs to slow the progression of kidney disease, particularly among participants with early stages of chronic kidney disease.

Added value of this study

Based on a joint analysis of 29 979 individuals in 41 treatment comparisons from randomised controlled trials, we showed that in studies of interventions in which albuminuria is considered to have biological plausibility as a surrogate endpoint, the treatment effect on albuminuria is significantly associated with the treatment effect on clinical endpoints. This association is stronger in populations with high baseline albuminuria (urine albumin-to-creatinine ratio >30 mg/g [3·4 mg/mmol]). Our model predicts that for future trials with sufficiently large sample sizes, treatments that decrease the geometric mean albuminuria relative to the control group to 0·7 (ie, 30% reduction in albuminuria) would have a high likelihood to confer clinical benefit.

Implications of all the available evidence

This study supports change in albuminuria as a surrogate endpoint for progression of chronic kidney disease in clinical trials, particularly in patients with high albuminuria. Future studies are warranted to determine how change in albuminuria can be best applied in designs for clinical trials of progression of chronic kidney disease while protecting against the risk of a false conclusion of clinical benefit.

randomised controlled trials in early stages of chronic kidney disease.

Change in albuminuria is a biologically plausible surrogate endpoint for the progression of chronic kidney disease in randomised controlled trials, and clinicians consider an early decrease in albuminuria to be indicative of a favourable response to treatment.[3–6] Albuminuria is a sensitive marker of the progression of kidney disease in early stages of chronic kidney disease and seems to be a cause of progression of chronic kidney disease in some kidney diseases.[3–5] In observational studies, increased albuminuria has consistently been shown to be strongly associated and a predictor of end-stage kidney disease.[7–12] However, uncertainty persists regarding the empirical evidence that treatment effects on change in albuminuria can be used to reliably predict treatment effects on the composite clinical endpoint (ie, end-stage kidney disease, doubling of serum creatinine, or eGFR of less than 15 mL/min per 1·73 m²).[13,14] Importantly, use of surrogate endpoints in clinical trials could decrease the number of participants and shorten the duration of follow-up required to achieve the statistical power to assess the efficacy of new interventions and assess interventions earlier in the disease course.

On March 15–16, 2018, the US National Kidney Foundation (NKF), in collaboration with the US Food and Drug Administration (FDA) and European Medicines Agency (EMA), sponsored a scientific workshop[15] to assess surrogate endpoints for trials of progression of kidney disease and improve understanding of change in albuminuria and GFR as measures of progression of kidney disease. As a result of this workshop, two meta-analyses were performed. Here we report the results of an individual patient-level Bayesian meta-analysis of treatment comparisons from randomised controlled trials to examine the agreement between treatment effects on early changes in albuminuria and treatment effects on the composite clinical endpoint for chronic kidney disease progression, and to inform the use of albuminuria as a surrogate endpoint in future randomised controlled trials. The companion report[16] provides the results of an observational analysis of 28 cohorts to examine the association between change in albuminuria and risk of end-stage kidney disease. Together, these analyses related to albuminuria provide a comprehensive assessment of the validity of using early changes in albuminuria as a surrogate endpoint for trials concerning the progression of chronic kidney disease.

Methods

Search strategy and selection criteria

Previously,[2,6] we did a systematic search of the MEDLINE database for randomised controlled trials published in English between Jan 1, 1946, and May 15, 2007. To update this dataset for the current analysis, we repeated the search for publications between May 16, 2007 (when the initial

search had been completed), and Dec 15, 2016, included the search terms "chronic kidney disease", "chronic renal insufficiency", "albuminuria", "proteinuria", "randomized controlled trial" (complete list of search terms is in the appendix). Inclusion criteria included that participants must be aged 18 years or older with an eGFR of more than 15 mL/min per 1·73 m², quantifiable measurements of albuminuria or proteinuria (ie, not dipstick) at baseline and within 12 months of follow-up, follow-up of participants for more than 12 months after baseline measurement, and information on the incidence of end-stage kidney disease (complete list of inclusion criteria is in the appendix). We had no exclusion criteria. Once eligible studies were identified, we contacted the authors of the studies to request access to the full participant-level data and those who agreed sent us data from their studies. The systematic search, data requests, and data analyses were done by members of the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) Data Coordinating Center. Disagreements about study eligibility and inclusion were discussed among the members of the CKD-EPI Data Coordinating Center with final decisions made by LAI, and disagreements over suitability of data resolved by contacting the study investigator to check if data were sufficient. Based on an a priori discussion with the NKF-FDA-EMA workshop Planning and Operations Committee, we excluded studies with interventions in which change in albuminuria was not thought to be biologically plausible as a surrogate endpoint (nurse-coordinated management and allopurinol; appendix).

This study was approved by Tufts Medical Center institutional review board. Informed consent was not required for the present analysis but was obtained for each individual randomised controlled trial.

Data analysis

Our primary objective was to assess the validity of change in albuminuria as a surrogate endpoint for progression of chronic kidney disease by modelling the association between treatment effects on 6-month change in urine albumin-to-creatinine ratio (ACR), and treatment effects on the composite clinical endpoint across the studies. Our secondary objective was to use these results to inform the use of change in ACR as a surrogate endpoint in future randomised controlled trials by estimating the probability of clinical benefit associated with a range of treatment effects on change in ACR.

We assessed the risk of bias for each randomised controlled trial using the risk-of-bias tool of the Cochrane collaboration (appendix).[17] For randomised controlled trials that investigated more than one intervention, we included a separate randomised treatment comparison for each independent treatment versus control comparison reported, so some participants were included in more than one analytical unit. We pooled randomised controlled trials with fewer than 100 participants that investigated the same disease and intervention.

Methods used to measure albuminuria varied between studies, with most measuring urine protein excretion rate (PER; appendix). Because guidelines[18] recommend use of urine ACR, we converted all urine PER measurements to ACRs using the equation ACR mg/g (mg/mmol) = 0·6 × PER mg per day. An early change in ACR was quantified as the change in log-transformed ACR from baseline to the measurement closest to 6 months (within 2·5 and 14 months) or 12 months (within 2·5 and 19 months).

See Online for appendix

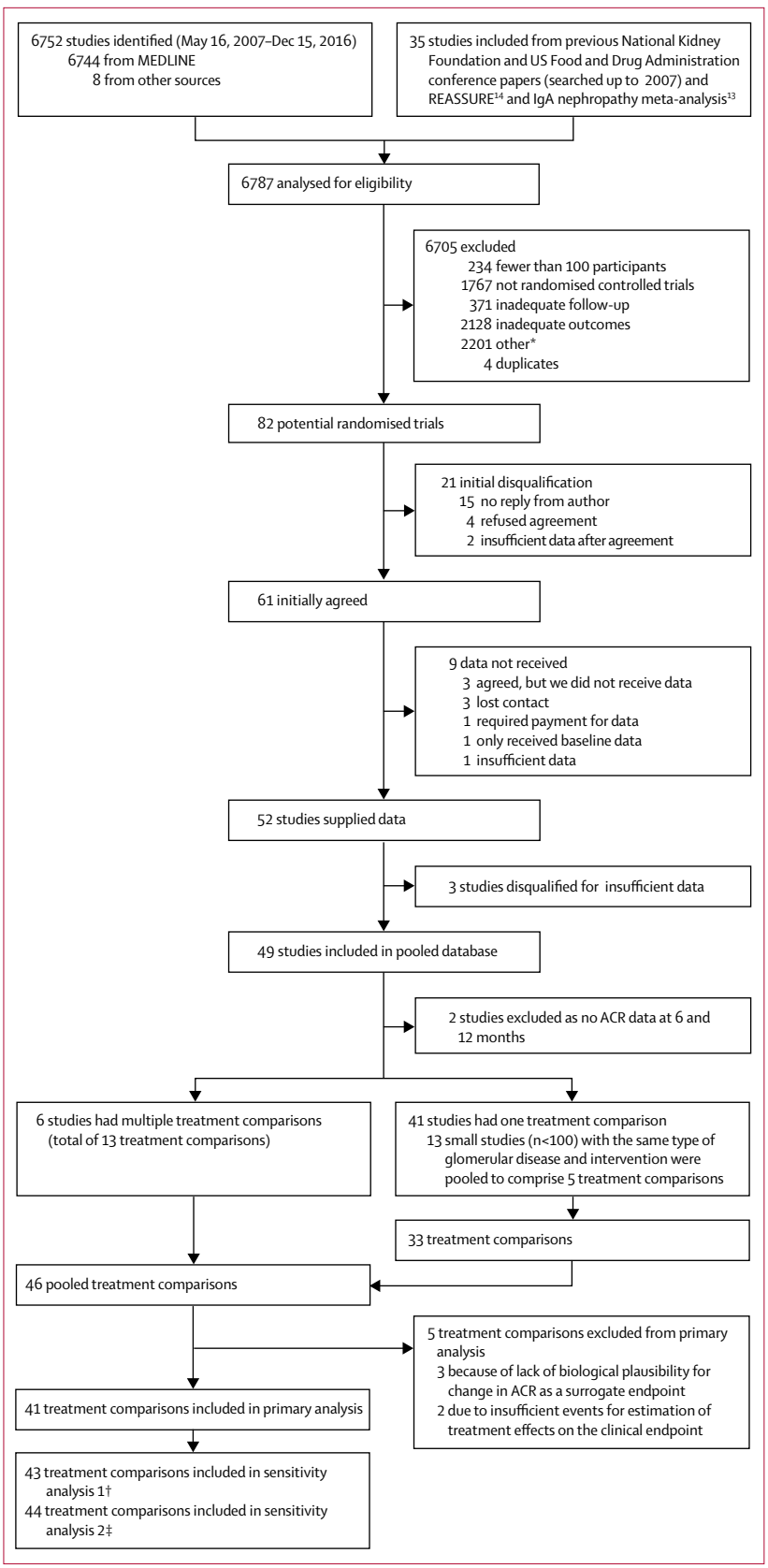
The clinical endpoint was defined as a composite of treated end-stage kidney disease (defined as initiation of chronic treatment with dialysis or kidney transplantation), eGFR of less than 15 mL/min per 1·73 m², or doubling of serum creatinine sustained at the subsequent visit.

For our primary analysis, we only used studies that had sufficient endpoints for estimation of treatment effects on the composite clinical endpoint. In sensitivity analyses, we used an alternative clinical endpoint, defined as end-stage kidney disease, eGFR of less than 15 mL/min per 1·73 m², or time to 40% decrease in eGFR.

Statistical analysis

We used descriptive statistics to report baseline demographic and clinical information of the participants in the eligible studies. Our trial-level analysis required two steps: first, intention-to-treat estimation of the treatment effects on the surrogate and clinical endpoints within each study, and second, a meta-regression to relate the treatment effects on the surrogate and clinical endpoints across studies. In the first step, treatment effects on change in ACR were estimated by analysing the covariance within each study, with the log-transformed change in ACR as the endpoint adjusting for treatment and log-transformed baseline ACR. We expressed treatment effects on ACR as geometric mean ratios. We estimated treatment effects on the clinical endpoint using Cox proportional hazard regressions to estimate hazard ratios (HRs) for the treatment in each study. We obtained summary estimates of treatment effects using a random-effects model, with heterogeneity between studies estimated with the I^2 statistic. In the second step, a Bayesian mixed-effects meta-regression analysis related the estimated treatment effects on the clinical endpoint to the estimated treatment effects on change in ACR, with the study as the unit of analysis (appendix). The model relates the treatment effects on the surrogate and clinical endpoints after accounting for random errors in the estimated effects in each randomised controlled trial.

The Bayesian meta-regression would support the validity of change in ACR as a surrogate endpoint if the following conditions were met: first, the slope of the meta-regression line is significant (ie, the Bayesian credible interval [BCI] of the slope does not cross zero); second, the intercept is close to zero, implying the absence of an average effect on the clinical endpoint when the treatment does not affect ACR; third, the



coefficient of determination (ie, R^2 value) is high, indicating a strong trial-level association and that treatment effects on ACR account for most of the variation in treatment effects on the clinical endpoint; and finally, the root mean square error is low, assuring low variation in the clinical endpoint given a fixed treatment effect on ACR. These metrics were expressed as median values and 2·5 to 97·5% Bayesian credible intervals (BCIs).

We used R^2 designations of trial-level associations of a weak association as R^2 less than 0·49, a moderate association as R^2 0·49–0·72, and a strong association as R^2 more than 0·72.[19]

We used positive predictive values to describe the uncertainty in predicting treatment effect on the clinical endpoint. From the trial-level meta-regression, we calculated 95% Bayesian prediction intervals and estimated the probabilities of clinical benefit (defined as HR <1) for an infinite, large, or modest sized randomised trial. A large randomised trial was defined as one in which the treatment effect on ACR can be estimated to within an SE of 0·05, corresponding to a sample size (n) of about 1090 with a minimum detectable geometric mean ratio of 0·849, and a modest sized randomised trial was defined as having an SE of 0·12 when n is roughly 190 with a minimum detectable geometric mean ratio of 0·675. We calculated the threshold associated with the smallest observed treatment effect on ACR that would assure a high probability of benefit of the treatment on the composite clinical endpoint, defined as a positive predictive value of 97·5%.

We did overall and subgroup trial-level analyses of the primary analytical dataset. Subgroups were defined by age (<60 years and ≥60 years), sex, race, baseline ACR (<30 mg/g and ≥30 mg/g), eGFR (<60 mL/min per 1·73m² and ≥60 mL/min per 1·73m²), disease (diabetes and diabetic kidney disease, glomerular diseases, or other unspecified causes of chronic kidney disease), and interventions. We categorised race in the subgroup analyses as black versus non-black individuals and calculated eGFR using the creatinine equation of the CKD-EPI.[20] Because of differences in the ranges of treatment effects, we determined that accuracy in predicting treatment effects on the clinical endpoint was best compared between subgroups using

Figure 1: **Study selection**
The primary analysis included studies with sufficient endpoints for estimation of the treatment effect on the clinical endpoint and interventions (referred to as studies) in which change in ACR is hypothesised to have biological plausibility as a surrogate. ACR=urine albumin-to-creatinine ratio.*Other reasons included no long-term kidney outcome, not studying progression of kidney disease, and not a primary investigation paper. †Sensitivity analysis 1 refers to the analysis that included treatment comparisons in studies with biological plausibility for change in ACR as a surrogate and sufficient alternative events. ‡Sensitivity analysis 2 refers to the analysis using only treatment comparisons in studies with sufficient endpoint events to estimate treatment effects on the clinical endpoint, regardless of biological plausibility.

	Studies	Participants	Age, years	Sex		Race or ethnicity		Diabetes	eGFR, mL/min per 1·73 m²	ACR, mg/g	Participants with clinical endpoint*	Interventions
				Female	Male	Black	Non-black					
Overall	41	29 979	58·2 (12·6)	9951 (33·2%)	20 028 (66·8%)	3833 (12·8%)	26 146 (87·2%)	21 206 (70·7%)	58·2 (25·0)	272 (30–1134)	3935 (13·1%)	NA
Diabetes	10	21 102	62·2 (9·9)	6527 (30·9%)	14 575 (69·1%)	1335 (6·3%)	19 767 (93·7%)	21 102 (100·0%)	61·4 (23·3)	270 (26–1126)	2103 (10·0%)	RAAS inhibitor *vs* calcium-channel blocker; low *vs* usual blood pressure; RAAS inhibitor *vs* control; sulodexide; and empagliflozin
Glomerular disease	9	1325	40·8 (12·9)	467 (35·2%)	858 (64·8%)	18 (1·4%)	1307 (98·6%)	5 (0·4%)	74·2 (29·7)	1311 (838–2335)	174 (13·1%)	Immunosuppression; and RAAS inhibitor *vs* control
Other chronic kidney disease	22	7552	50·1 (12·9)	2957 (39·2%)	4595 (60·8%)	2480 (32·8%)	5072 (67·2%)	99 (1·3%)	46·6 (24·5)	126 (30–838)	1658 (22·0%)	RAAS inhibitor *vs* control; RAAS inhibitor *vs* calcium-channel blocker; low *vs* usual blood pressure; albuminuria-targeted protocol; and low *vs* usual protein diet

Data are n, mean (SD), n (%), or median (IQR). Other chronic kidney disease refers to causes of chronic kidney disease other than glomerular disease or diabetes, or if cause was not specified. Participants with clinical endpoint defined as the composite of participants who had chronic dialysis or kidney transplantation, eGFR less than 15 mL/min per 1·73 m², or confirmed doubling of serum creatinine. eGFR=estimated glomerular filtration rate. ACR=urine albumin-to-creatinine ratio. NA=not applicable. RAAS=renin–angiotensin–aldosterone system.

Table 1: **Clinical characteristics of the population, by disease aetiology**

the root mean square error. In sensitivity analyses, we repeated the analyses in the full dataset (ie, not restricted to studies in which change in ACR is hypothesised to have biological plausibility as a surrogate), using change in ACR at 12 months, the alternative clinical endpoint, and using a clinical endpoint that also includes death, and excluding studies determined to be outliers or with little variability in the serum creatinine measurements used to define the clinical endpoint

We did analyses using SAS version 9.4 and R version 3.16.1.

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

We identified 6752 studies through our literature search and an additional 35 from other sources; 82 (1%) were identified as potentially eligible randomised controlled trials after removal of duplicates. 49 (60%) of these 82 studies were included in the pooled database after disqualification of studies for non-agreement to participate (n=21), not receiving data after agreement (n=9), and insufficient data (n=3). Of the pooled studies, 46 (94%) had treatment comparisons for albuminuria, of which three (7%) were excluded because albuminuria was not hypothesised to be biologically plausible as a surrogate endpoint (figure 1). Hence, 43 treatment comparisons from clinical trials (referred to as studies throughout), comprising 30 078 participants with interventions for which the mechanisms were hypothesised to affect albuminuria were included in the full dataset. Two studies in the full dataset were excluded from the primary analysis because they did not have sufficient endpoints for estimation of treatment effects on the clinical endpoint, leaving 41 studies with 29 979 participants, of whom 21 206 (71%) had diabetes for the primary analysis.

For the primary analysis dataset, aggregate characteristics of the study populations included are in table 1 (stratified by overall disease) and the appendix (stratified by sex; furthermore, characteristics by study are in the appendix). The mean age of study participants ranged from 31·6 years (SD 11·5) to 71·4 years (SD 8·6) and the proportion of male participants ranged from 16% to 90%. Across the disease groups, the average baseline eGFR was 58·2 mL/min per 1·73 m² (SD 25·0) and the median baseline ACR was 272 mg/g (IQR 30–1134).

Over a median period of 6 months (IQR 6–6·4), the geometric mean ratio of ACR in the control groups was 16% (95% CI 8–24) and in the treatment groups was 34% (27–40), resulting in a 22% decrease in ACR due to treatment (geometric mean ratio 0·78, 95% CI 0·74–0·82; figure 2; appendix), with larger effects at baseline ACR levels greater than or equal to 30 mg/g (3·4 mg/mmol) compared with ACR <30 mg/g; geometric mean ratio 0·76, 95% CI 0·72–0·81, *vs* 0·92, 0·89–0·96). Similar results were seen for subgroups defined by age, sex, race (black *vs* non-black individuals), baseline eGFR, and cause of disease, and when the change in ACR was analysed over 12 months (appendix).

For the primary analysis dataset, over a median follow-up of 3·4 years (IQR 2·3–4·2), 3935 (13·1%) of 29 979 participants reached the composite clinical

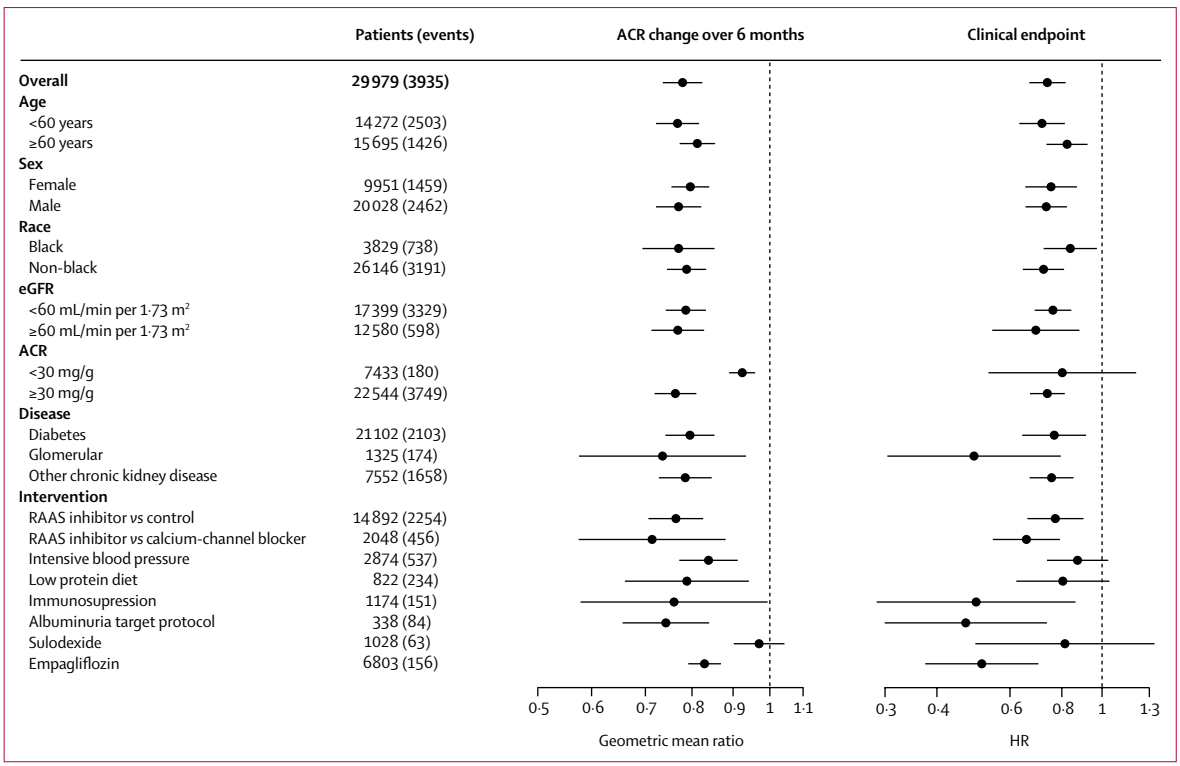


Figure 2: **Treatment effect on change in albuminuria at 6 months and on the clinical endpoint in the overall population and subgroups**
Data points are the estimated treatment effect and horizontal lines the 95% CIs. Data for all studies are in the appendix. Treatment effect on albuminuria is expressed as the geometric mean ratio of ACR; to convert geometric mean ratio to percentage decrease in ACR: (1–geometric mean ratio) × –100. Treatment effect on the clinical endpoint is expressed as an HR. The composite clinical endpoint was end-stage kidney disease, doubling of serum creatine concentration, or eGFR of less than 15 mL/min per 1·73 m². ACR was log transformed in each analysis. Other chronic kidney disease refers to causes of chronic kidney disease other than glomerular disease or diabetes, or that the cause was not specified. ACR=albumin to creatinine ratio. eGFR=estimated glomerular filtration rate. HR=hazard ratio. RAAS=renin–angiotensin–aldosterone system

endpoint (appendix). Across all interventions, the active treatment led to a decrease in risk of the clinical endpoint (HR 0·73, 95% CI 0·67–0·81), with similar results across the subgroups (figure 2; appendix). The sensitivity analysis was consistent with these results (appendix).

Results of the trial-level analysis of the association between the treatment effects on change in ACR and the treatment effects on the clinical endpoint are shown in figure 3 and the appendix. The slope on the log scale was 0·89 (95% BCI 0·13 to 1·70), which indicates that, for example, a 30% increase in the treatment effect on the geometric mean ACR would be associated with an average 27% decrease in the hazard for the clinical endpoint (95% BCI 5 to 45). The intercept of the regression line was –0·07 (95% BCI –0·29 to 0·14), indicating that for no change in ACR, the probability of a large treatment effect on the clinical endpoint was low. The median estimate for R^2 was 0·47 (95% BCI 0·02 to 0·96) and the root mean square error was 0·14 (95% BCI 0·03 to 0·27). Reflecting the wide BCI of the R^2, the trial-level association has Bayesian probabilities of 53% weak association, 28% moderate association, and 19% strong association ranges for the strength of a surrogate endpoint. The association strengthened when the analyses were restricted to the subgroup of patients with a baseline ACR of more than 30 mg/g (R^2 0·72, 95% BCI 0·05–0·99; figure 3; appendix), with trial-level association Bayesian probabilities of 27% for a weak association, 24% for a moderate association, and 49% for a strong association.

For the trial-level analysis, overlapping BCIs indicated no clear evidence of significant differences in the root mean square error were found among the subgroups stratified by eGFR or cause of disease, but the BCIs were wide and so the precision of the results is not sufficient to be conclusive (table 2). Results were consistent with the overall analysis when renin–angiotensin–aldosterone system (RAAS) blockade interventions versus control treatments were analysed separately (median R^2 0·64, 95% BCI 0·00–0·99; root mean square error 0·11, 95% BCI 0·02–0·33; appendix). We had similar results for the treatment effects on ACR when they were analysed at 12 months, and for the alternative clinical endpoint and when death was included as part of the composite with the clinical endpoint (appendix), but when the full dataset of 43 studies was analysed the association was slightly weaker with wider confidence bands than when only the primary dataset of 41 studies

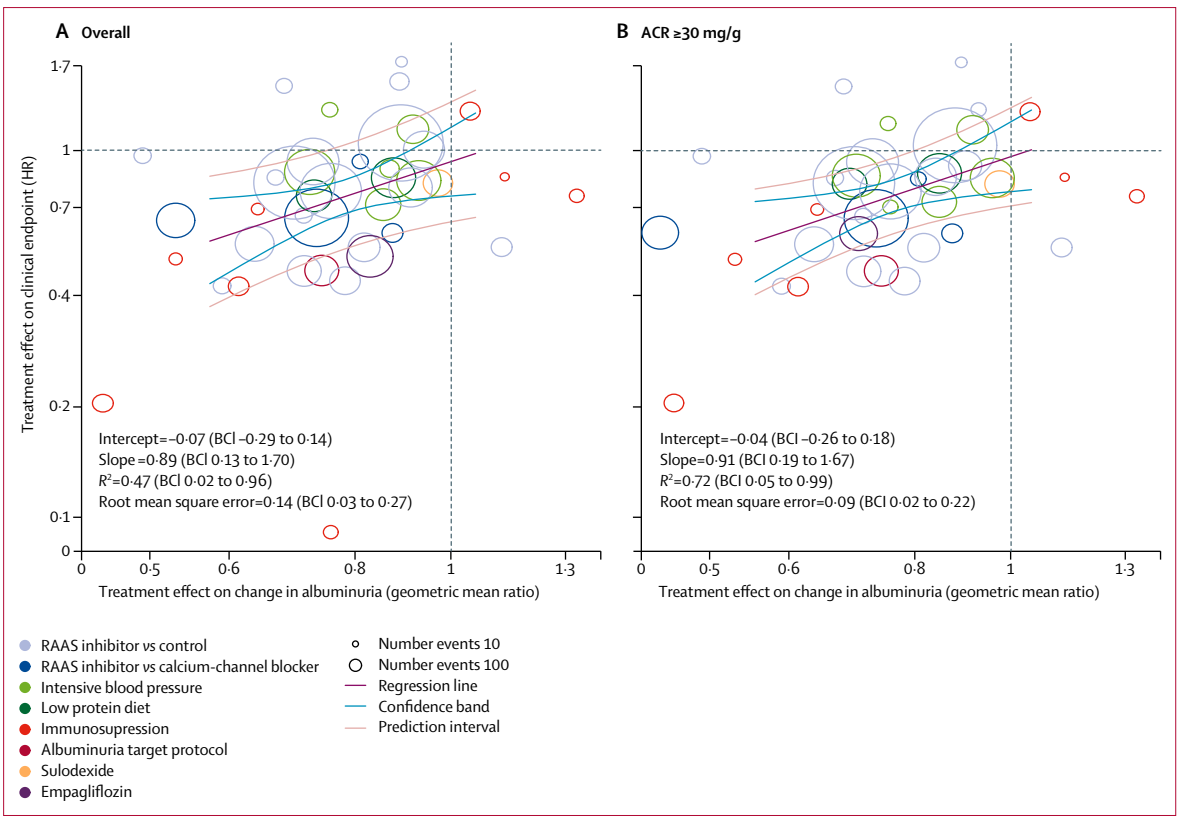


Figure 3: **Trial-level analyses for the association between treatment effects on change in albuminuria and treatment effects on the clinical endpoint for the pooled population (A) and for participants who had baseline ACR of more than 30 mg/g (B)**
The vertical axes are the estimated treatment effects on the clinical endpoint (HR) and the horizontal axes are the estimated treatment effects on the 6 month change in albuminuria (geometric mean ratio of log-transformed ACR). The composite clinical endpoint was end-stage kidney disease, doubling of serum creatine concentration, or eGFR of less than 15 mL/min per 1·73 m². The different coloured circles indicate intervention types, and each circle is a separate intervention with the size of the circle proportional to the number of events. The line of regression through the studies, Bayesian confidence bands, and Bayesian prediction bands from the model are shown. ACR=urine albumin-to-creatinine ratio. BCI=Bayesian credible interval. eGFR=estimated glomerular filtration rate. HR=hazard ratio. RAAS=renin–angiotensin–aldosterone system. eGFR=estimated glomerular filtration rate.

was used (appendix). Notably, among participants who had a baseline ACR of less than 30 mg/g the event rate was low (180 [4·6%] of 3935 events), and so the results for this subgroup are unlikely to be precise.

We identified the EMPA-REG OUTCOME study[21] as an outlier because of a weaker association between the treatment effects on the surrogate endpoint and the treatment effects on the clinical endpoint compared with other studies. We therefore repeated our analysis of the primary dataset excluding this study and found the strength of the association increased (from R^2 0·47, 95% BCI 0·02–0·96, to R^2 0·72, 0·08–0·99). However, we did not see this effect when we restricted the analysis to the subgroup of participants with baseline ACR levels of more than 30 mg/g; R^2 0·72 (95% BCI 0·05–0·99) with inclusion of EMPA-REG OUTCOME, and R^2 0·68 (0·04–0·99) after exclusion (appendix). We also identified the study by Hou and colleagues[22] as having unusually small variability in the serum creatinine measurements used to define the clinical endpoints, and so we repeated our primary analysis excluding this study and found our results were consistent whether this study was included or excluded (R^2 0·47 [95% BCI 0·02–0·96] with inclusion of Hou and colleagues' study, and R^2 0·42 [0·01–0·95] after exclusion).

For application of ACR as a surrogate endpoint in future randomised controlled trials, table 3 shows the predicted HRs and 95% Bayesian prediction intervals for the treatment effect on the clinical endpoint and the corresponding positive predictive values across a range of magnitudes of ACR treatment effects, overall and for participants with baseline ACR levels of more than 30 mg/g. For a study of infinite or large sample size, treatments that decrease the geometric mean albuminuria by 30% (ie, ACR geometric mean ratio 0·7) compared with the control group will provide an average HR for the clinical endpoint of 0·68 (95% prediction interval 0·47–0·95). A treatment effect of 25% decrease in ACR would be required to provide 97·5% confidence of a clinical benefit, with a 31% decrease required for modest sized trials. If participants were restricted to those with a baseline ACR level of more than 30 mg/g, a 20% decrease

	Studies (interventions)	Patients (events)	Slope	Intercept	R^2	Root mean square error
Overall	41 (8)	29 979 (3935)	0·89 (0·13 to 1·70)	−0·07 (−0·29 to 0·14)	0·47 (0·02 to 0·96)	0·14 (0·03 to 0·27)
eGFR						
<60 mL/min per 1·73m²	39 (8)	17 387 (3329)	0·89 (0·04 to 1·83)	−0·03 (−0·27 to 0·22)	0·62 (0·01 to 0·99)	0·09 (0·02 to 0·23)
≥60 mL/min per 1·73 m²	23 (6)	12 348 (598)	2·15 (−1·49 to 7·52)	0·13 (−0·71 to 1·23)	0·77 (0·01 to 1·00)	0·14 (0·02 to 0·50)
ACR						
<30 mg/g*	10 (5)	7401 (180)	−9·86 (−53·39 to 45·19)	−1·07 (−5·01 to 3·39)	0·96 (0·02 to 1·00)	0·07 (0·01 to 0·60)
≥30 mg/g	41 (8)	22 544 (3749)	0·91 (0·19 to 1·67)	−0·04 (−0·26 to 0·18)	0·72 (0·05 to 0·99)	0·09 (0·02 to 0·22)
Disease						
Diabetes	10 (5)	21 102 (2103)	0·41 (−2·10 to 2·67)	−0·16 (−0·78 to 0·39)	0·13 (0·00 to 0·86)	0·20 (0·04 to 0·47)
Glomerular disease	9 (2)	1352 (174)	1·63 (0·19 to 3·95)	−0·16 (−0·77 to 0·68)	0·98 (0·11 to 1·00)	0·06 (0·01 to 0·57)
Other chronic kidney disease	22 (5)	7552 (1658)	0·73 (−0·16 to 1·76)	−0·10 (−0·34 to 0·17)	0·75 (0·01 to 0·99)	0·05 (0·01 to 0·22)
Disease when ACR ≥30 mg/g						
Diabetes	10 (5)	15 532 (2030)	1·10 (−0·76 to 2·72)	0·06 (−0·45 to 0·48)	0·63 (0·00 to 0·99)	0·08 (0·02 to 0·32)
Glomerular disease	9 (2)	1324 (174)	1·63 (0·12 to 3·91)	−0·16 (−0·78 to 0·65)	0·98 (0·11 to 1·00)	0·06 (0·01 to 0·56)
Other chronic kidney disease	22 (5)	5688 (1545)	0·53 (−0·38 to 1·53)	−0·15 (−0·42 to 0·14)	0·65 (0·00 to 0·99)	0·05 (0·01 to 0·21)

Data are n (n), or median (2·5 to 97·5 Bayesian credible intervals). Other chronic kidney disease refers to causes of chronic kidney disease other than glomerular disease or diabetes, or for which cause was not specified. eGFR=estimated glomerular filtration rate. ACR=albumin to creatinine ratio. *Event rate <5% so estimates are unreliable.

Table 2: **Trial-level analysis by subgroups**

	Infinite sample size		Large trial		Modest trial	
	HR	PPV	HR	PPV	HR	PPV
All participants						
0·5	0·50 (0·30–0·80)	1·00	0·50 (0·30–0·80)	1·00	0·51 (0·28–0·82)	1·00
0·6	0·59 (0·39–0·86)	0·99	0·59 (0·39–0·87)	0·99	0·60 (0·36–0·90)	0·99
0·7	0·68 (0·47–0·95)	0·98	0·68 (0·46–0·96)	0·98	0·68 (0·44–1·01)	0·97
0·8	0·77 (0·53–1·06)	0·96	0·76 (0·53–1·07)	0·95	0·76 (0·50–1·14)	0·92
0·9	0·85 (0·59–1·19)	0·86	0·85 (0·58–1·21)	0·84	0·84 (0·56–1·30)	0·81
1·0	0·93 (0·63–1·35)	0·66	0·93 (0·62–1·37)	0·66	0·92 (0·60–1·48)	0·65
Threshold to ensure PPV ≥97·5%	0·75	..	0·74	..	0·69	..
Participants with baseline ACR ≥30 mg/g						
0·5	0·51 (0·34–0·76)	1·00	0·52 (0·33–0·76)	1·00	0·52 (0·32–0·77)	1·00
0·6	0·61 (0·44–0·81)	1·00	0·61 (0·43–0·81)	1·00	0·61 (0·41–0·84)	1·00
0·7	0·70 (0·53–0·89)	0·99	0·70 (0·52–0·89)	0·99	0·70 (0·49–0·96)	0·98
0·8	0·79 (0·60–0·99)	0·98	0·79 (0·59–1·02)	0·97	0·79 (0·56–1·10)	0·93
0·9	0·88 (0·65–1·13)	0·87	0·88 (0·64–1·16)	0·85	0·87 (0·62–1·26)	0·80
1·0	0·97 (0·69–1·30)	0·60	0·96 (0·69–1·34)	0·61	0·95 (0·66–1·45)	0·60
Threshold to ensure PPV ≥97·5%	0·80	..	0·79	..	0·73	..

Treatment effect on ACR is expressed as geometric mean ratio and change in albuminuria is expressed over 6 months. Treatment effect on the clinical endpoint is expressed as a hazard ratio (median, 95% Bayesian prediction intervals). Percentage decrease in ACR = (1−geometric mean ratio) × −100. The composite clinical endpoint was end-stage kidney disease, doubling of serum creatine concentration, or an estimate glomerular filtration rate of less than 15 mL/min per 1·73 m². A modest trial was defined as one that results in treatment effect of albuminuria with SE of 0·12, minimal detectable geometric mean ratio of 0·675 and approximate sample size of 190, and large trial was defined as one with SE of 0·05, minimal detectable geometric mean ratio of 0·849 and approximate sample size of 1090. ACR=urine albumin-to-creatinine ratio. HR=hazard ratio. PPV=positive predictive value.

Table 3: **Predicted treatment effect of change in ACR on the clinical endpoint and positive predictive value, by size of randomised controlled trial**

in ACR would be required for studies of infinite sample size, a 21% decrease for trials with a large sample size, and a 27% decrease for trials of modest size. We saw similar results for changes in albuminuria at 12 months (appendix).

Discussion

This study provides a comprehensive assessment of change in albuminuria as a surrogate endpoint on the basis of a joint analysis of 29 979 individuals in 41 studies. We found that in studies of interventions in which

albuminuria is considered to have biological plausibility as a surrogate endpoint, the treatment effect on albuminuria had moderately strong associations with treatment effect on clinical endpoints. The association was stronger for participants with baseline ACR levels of more than 30 mg/g (equivalent to 3·4 mg/mmol), whereas we found no evidence for differing effects by baseline level of eGFR or for other subgroups (ie, age, sex, baseline ACR, baseline eGFR, disease, or interventions). We provide thresholds for minimum effects on change in albuminuria that provide high confidence for significant treatment effects on the clinical endpoint, providing guidance about how to implement change in ACR as an endpoint in future randomised controlled trials. The companion meta-analysis of observational studies[16] examined the association between changes in ACR with subsequent risk of end-stage kidney disease in approximately 700 000 individuals and showed that, after adjustment for measurement error, a 30% decrease in albuminuria assessed over a period of 2 years was associated with a 22% reduction in the risk of end-stage kidney disease (HR 0·78, 95% CI 0·66–0·92). The similarity between the results from that report and ours provides reassurance about the robustness of the findings and the utility of changes in albuminuria to predict clinical prognosis.[16] Together, these results support the use of early change in albuminuria as a surrogate endpoint for randomised controlled trials of chronic kidney disease progression.[10–14,23]

Previous meta-analyses reported conflicting results about the validity of change in albuminuria as a surrogate endpoint.[14,23,24] One large meta-analysis on evaluating surrogate endpoints,[13] involving 9088 participants in 37 trials, showed no clear association between early changes in albuminuria and the clinical endpoint, but it had insufficient variation in treatment effects or statistical power, or both, to adequately address the question.[13] Another meta-analysis[14] involved larger clinical trials and reported a strong association between treatment effects on early and late changes in albuminuria and end-stage kidney disease, which was consistent across drug classes and patient characteristics. However, that study did not account for correlation among the sampling errors in the treatment effects, as we did through our Bayesian meta-regression.[4] A more recent meta-analysis[24] did not find a significant association between changes in albuminuria and end-stage kidney disease, but was limited by an insufficient number of studies and range of interventions, and a lack of harmonisation of the endpoints.[25] The different conclusions from these previous meta-analyses probably result from differences in the included trials and analytical methodology, precluding definitive conclusions regarding the validity of change in albuminuria as a surrogate endpoint. Our study advances these previous analyses. To our knowledge, this study is the largest analysis to address this topic to date, with the most diverse dataset of included trials, which is necessary for a robust trial-level analysis. The Bayesian analyses also allowed us to articulate thresholds for treatment effect on change in albuminuria that are sufficiently large to result in high confidence that the treatment effect on the clinical endpoint would also be significant. Similar analyses using less formal analyses with fewer trials than ours have been done to establish blood pressure and plasma cholesterol as surrogates for cardiovascular endpoints.[26–28]

The stronger association between treatment effects on change in albuminuria and treatment effects on the clinical endpoint in subgroups with higher baseline albuminuria is consistent with the hypothesis that the mechanism of the beneficial effect of interventions is through the lowering of albuminuria and implies that future randomised controlled trials that use change in albuminuria as an endpoint should be restricted to participants who have some degree of albuminuria. Indeed, one study, EMPA-REG OUTCOME,[21] appeared to be an outlier only because its average level of albuminuria was substantially lower than most other included studies. Similarly, the stronger associations found when we restricted the analyses to interventions in which change in albuminuria was considered to have biological plausibility as a surrogate endpoint suggests that albuminuria cannot be used independently of the hypothesised mechanism of the effect of the intervention on albuminuria. We did not have sufficient power to make definitive conclusions for or against differences in trial-level associations by the baseline GFR level and aetiology of chronic kidney disease. Nevertheless, in the absence of an a priori hypothesis that results would differ among these groups, together with the similarity of the results for treatment effects on albuminuria and on the clinical endpoint by GFR level and disease, our interpretation is that these results provide support for the use of change in albuminuria as a surrogate endpoint across the range of GFR and disease aetiologies investigated here.

The requirement for changes in albuminuria of about 20–30% for large sample sizes in future trials might appear to be excessively severe and not obtainable in many settings; however, these thresholds increase the probability of a significant treatment effect on the clinical endpoint, essentially making albuminuria a stronger surrogate. Additionally, use of change in albuminuria as a surrogate endpoint shortens the required duration of the trial, which can lead to more efficient and probably less expensive study designs. Increasing the number of albuminuria measurements will most likely increase the precision of the treatment effect[29] and decrease the required threshold for change in albuminuria when the sample size is modest. In our view, these results are particularly useful for diabetic kidney disease with high albuminuria at baseline. Diabetic kidney disease is highly prevalent and is associated with a high risk of progression to kidney failure, and few therapies are available, making it a condition with substantial unmet clinical need. Additionally, our results are complementary to ongoing

work to define clinical endpoints for randomised trials in glomerular diseases with nephrotic syndrome or IgA nephropathy.[30,31] We do not suggest that our findings should replace these initiatives; indeed, the large sample size needed to reliably assess treatment effects on albuminuria would not be possible in such rare diseases. Nevertheless, investigators designing studies in rare diseases could still use our results for positive predictive values but apply less stringent criteria for a minimal treatment effect on change in albuminuria. These results might also be applicable to randomised trials investigating interventions in patients with chronic kidney disease of other or unspecified cause in whom absolute risk is high (eg, the subgroup with high baseline levels of albuminuria). Crucially, shorter trials based on changes in albuminuria would require additional longer-term assessment of safety.

These analyses had several limitations. First, although we used an unbiased systematic review and obtained, to our knowledge, the most randomised controlled trials used in any assessment of a surrogate endpoint in chronic kidney disease, the number of studies with available data was not very large, and some trialists declined to participate. As such, some imprecision exists in our estimates of trial-level associations, particularly within subgroups, and our analysis had insufficient power to adequately determine differences in the results by the cause of chronic kidney disease, race, or ethnicity. Notably, the proportion of women included was lower than the proportion of men, and although our results provide no evidence that the trial-level association differs between men and women, our analyses did not have sufficient power to rule out such a difference. Furthermore, among the studies included, categorisation of race and ethnicity was inconsistent, and so we could not comprehensively categorise participants by their race and ethnicity. Second, we only included clinical endpoints that occurred during the course of the trial. As such, disagreement between the surrogate and clinical endpoint could arise from insufficient follow-up to capture the longer-term treatment effect on the clinical endpoint. Third, the observed heterogeneity in treatment effects across the trials indicate not only biological differences between treatments but also heterogeneity in the design and procedures between clinical trials, thereby diluting the strength of the true association. Fourth, our prediction intervals and positive predictive values at different magnitudes of treatment effects on change in ACR show not only the implications of varying treatment effects on ACR, but also the specific results in previous studies, which showed an average beneficial effect on the clinical endpoint. As a result, achieving a treatment effect on change in ACR large enough to provide a high positive predictive value does not necessarily guarantee a low risk of a false conclusion of clinical benefit for an ineffective treatment. Additional work is required to determine how albuminuria can be used in the design of clinical trials to protect against the risk of false-positive conclusions under the null hypothesis of no clinical benefit. Finally, most studies included in these analyses tested RAAS blockers. The strength of the association was similar when we assessed only trials comparing these drugs versus control, but insufficient data were available to analyse other interventions separately.

When interpreted in conjunction with experimental evidence and findings from observational studies, including the companion meta-analysis of observational studies,[16] the results of our trial-level analyses support a role for early change in albuminuria as a surrogate endpoint for progression of chronic kidney disease in clinical trials, particularly in trials with entry criteria that restrict enrolment to patients with high baseline albuminuria. However, the use of change in albuminuria as a surrogate endpoint does have limitations: particularly, inferring the clinical benefit of reduction in albuminuria does not seem to be possible for interventions that result in moderate decreases in albuminuria (ie, <20–30%). Future research is needed to elucidate how change in albuminuria can be optimally applied in the design of clinical trials investigating chronic kidney disease while protecting against the risk of a false conclusion of clinical benefit.

Contributors

HJLH, TG, JC, ASL, RTG, and LAI conceived the study concept and design. The Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) investigators and collaborators acquired the data. HT, TG, ALS, and LAI analysed the data. All authors contributed to the interpretation of the data. HJLH and LAI drafted the report and all authors provided critical revisions for important intellectual content. All collaborators shared data and were given the opportunity to comment on the report. HJLH, TG, ASL, and LAI obtained funding for CKD-EPI; individual cohort and collaborator support is listed in the appendix. The Planning and Operations Committee of the collaborative US National Kidney Foundation, US Food and Drug Administration, and European Medicines Agency workshop contributed to the design and critical review of these analyses.

Chronic Kidney Disease Epidemiology Collaboration

Investigators and collaborators: African American Study of Kidney Disease and Hypertension (AASK)—Tom Greene; Appropriate Blood Pressure Control in Diabetes trial (ABCD)—Robert W Schrier, Raymond O Estacio; Action in Diabetes and Vascular Disease: Preterax and Diamicron Modified Realease Controlled Evaluation (ADVANCE) trial—Vlado Perkovic; ACE Inhibition in Progressive Renal Insufficiency (AIPRI)—Giuseppe Maschio, Francesco Locatelli; Aliskiren Trial in Type 2 Diabetes Using Cardiorenal Endpoints (ALTITUDE)—Hans-Henrik Parving; Randomized controlled clinical trial in proteinuric IgA nephropathy (Bari, Italy)—Francesco Paolo Schena; Manno Carlo;Bologna (Italy)—Pietro Zucchelli; Boston(MA, USA)—Barry M Brenner; Canadian Prevention of Renal and Cardiovascular Endpoints Trial (CanPREVENT)—Brendan Barret; Copenhagen(Denmark)—Anne-Lise Kamper, Svend Strandgaard; Collaborative Study Group (CSG)—Roger A Rodby, Richard D Rohde, Julia B Lewis, Edmund Lewis; Empagliflozin Cardiovascular Outcome Event Trial in Type 2 Diabetes Mellitus Patients (EMPA-REG OUTCOME)—Christoph Wanner, Maximilian von Eynatten; Fukuoka (Japan)—Ritsuko Katafuchi; Groningen (The Netherlands)—Paul E de Jong, G G van Essen; Guangzhou (China)—Fan Fan Hou, Di Xie; Halt Progression of Polycystic Kidney Disease study (HALT-PKD)—Ronald D Perrone, Kaleab Z Abebe; Hong Kong Lupus Nephritis (HKLN)—Tak Mao Chan; Hong Kong study using Valsartan in IgA Nephropathy (HKVIN)—Philip Li, C B Leung, C C Szeto, K M Chow; Irbesartan Diabetic Nephropathy Trial (IDNT)—

Edmund Lewis, Lawrence G Hunsicker, Julia B Lewis; Lecco (Italy)—Francesco Locatelli, Lucia Del Vecchio, Simeone Andrulli, Claudio Pozzi; Leuven (Belgium)—Bart Maes; Lupus Nephritis Collaborative Study (LNCS)—Julia B Lewis, Jamie Dwyer, Edmund Lewis, John M Lachin; Madrid (Spain)—Marian Goicoechea, Eduardo Verde, Ursula Verdalles Jose Luño; Tacrolimus in membranous nephropathy (Madrid, Spain)—Manuel Praga, Fernando Caravaca, Eduardo Gutierrez, Angel Sevillano; Multifactorial Approach and Superior Treatment Efficacy in Renal Patients with the Aid of Nurse Practitioners study (MASTERPLAN)—Jack F M Wetzels, Peter J Blankestijn, Arjan D van Zuilen, Jan van den Brand; Modification of Diet in Renal Disease study (MDRD Study)—Gerald Beck, Tom Greene, John Kusek, Saulo Klahr; Milan (Italy)—Claudio Ponticelli Montagnino, Patrizia Passerini, Gabriella Moroni, Giuseppe Montogrino; New York (NY, USA)—Gerald B Appel, Gershon Frisch; Olmesartan Reducing Incidence of Endstage Renal Disease in Diabetic Nephropathy Trial (ORIENT)—Fumiaki Kobayashi, Hirofumi Makino, Sadayoshi Ito, Enyu Imai; Ramipril Efficacy In Nephropathy study 1 (REIN 1)—Giuseppe Remuzzi, Piero Ruggenenti; Ramipril Efficacy In Nephropathy study 2 (REIN 2)—Giuseppe Remuzzi, Piero Ruggenenti; Reduction of Endpoints in Non-insulin Dependent Diabetes Mellitus with the Angiotensin II Antagonist Losartan (RENAAL)—Dick De Zeeuw, Hiddo J Lambers Heerspink, Barry M Brenner, William Keane; Renoprotection of Optimal Antiproteinuric Doses (ROAD) study—Fan Fan Hou; Rochester (NY, USA)—James Donadio, Fernando C Fervenza; Supportive Versus Immunosuppressive Therapy for the Treatment of Progressive IgA Nephropathy trial (STOP-IgAN)—Jürgen Floege, Thomas Rauen, Christina Fitzner, Ralf-Dieter Hilgers; Strasbourg (France)—Thierry P Hannedouche; Sulodexide Macroalbuminuria trial (SUN-MACRO)—Julia B Lewis, Jamie Dwyer, Edmund Lewis; Dallas (TX, USA)—Robert D Toto; and Parkville (VIC, Australia)—Gavin J Becker, Benno U Ihle, Priscilla S Kincaid-Smith.

Planning and Operations Committee for the US National Kidney Foundation, US Food and Drug Administration, and European Medicines Agency workshop

Andrew Levey (chair), Ron T Gansevoort, Josef Coresh, Dick de Zeeuw, Kai-Uwe Eckardt, Hrefna Gudmundsdottir, Adeera Levin, Romaldas Maciulaitis, Tom Manley, Vlado Perkovic, Kimberly Smith, Norman Stockbridge, Aliza Thompson, Thorsten Vetter, Kerry Willis, Luxia Zhang.

Declaration of interests

HJLH reports grants and other support from AbbVie, AstraZeneca, Boehringer Ingelheim, and Janssen; and other support from Astellas, Fresenius, Gilead, and Merck outside of the submitted work. TG reports grants from the US National Kidney Foundation (NKF) during the conduct of the study; and personal fees from DURECT Corportation, Janssen, and Pfizer outside of the submitted work. JC reports grants from the US National Insitutes of Health (NIH) and NKF during the conduct of the study. FFH reports personal fees from AbbVie and AstraZeneca outside of the submitted work. FL reports personal fees from Amgen, Roche, Akebia, Astellas, GlaxoSmithKline, AbbVie, Vifor Fresenius Pharma, and Menarini outside of the submitted work. ASL reports grants from NKF and NIH during the conduct of the study; and other support from Siemens outside of the submitted work. LAI reports grants from NIH and NKF during the conduct of the study and outside of the submitted work; funding from NIH, NKF, Retrophin, Omeros and Reata Pharmaceuticals for research and contractswith Tufts Medical Center; and consulting agreements with Tricida and Omeros Corp. ASL, JC, and LAI also report a provisional patent (filed Aug 15, 2014; precise estimation of glomerular filtration rate from multiple biomarkers, patent number PCT/US2015/044567), and Tufts Medical Center, John Hopkins University, and Metabolon have a collaboration agreement to develop a product to estimate GFR from a panel of markers. All other authors declare no competing interests.

Acknowledgments

The study was funded by the US National Kidney Foundation. We thank Juhi Chaudhari and Riley Marotta for help with preparation of the manuscript and Shiyuan Miao for help with preparation of the figures. A variety of sources have supported the randomised controlled trials included in the meta-analysis. These funding sources include government agencies such as national institutes of health and medical research councils, as well as foundations and industry sponsors (appendix).

References

1 Levey AS, Inker LA, Matsushita K, et al. GFR decline as an end point for clinical trials in CKD: a scientific workshop sponsored by the National Kidney Foundation and the US Food and Drug Administration. *Am J Kidney Dis* 2014; **64:** 821–35.
2 Schievink B, Kröpelin T, Mulder S, et al. Early renin–angiotensin system intervention is more beneficial than late intervention in delaying end-stage renal disease in patients with type 2 diabetes. *Diabetes Obes Metab* 2016; **18:** 64–71.
3 Burton C, Harris KP. The role of proteinuria in the progression of chronic renal failure. *Am J Kidney Dis* 1996; **27:** 765–75.
4 Abbate M, Zoja C, Remuzzi G. How does proteinuria cause progressive renal damage? *J Am Soc Nephrol* 2006; **17:** 2974–84.
5 Tryggvason K, Pettersson E. Causes and consequences of proteinuria: the kidney filtration barrier and progressive renal failure. *J Intern Med* 2003; **254:** 216–24.
6 Ruggenenti P, Perticucci E, Cravedi P, et al. Role of remission clinics in the longitudinal treatment of CKD. *J Am Soc Nephrol* 2008; **19:** 1213–24.
7 Carrero JJ, Grams ME, Sang Y, et al. Albuminuria changes are associated with subsequent risk of end-stage renal disease and mortality. *Kidney Int* 2017; **91:** 244–51.
8 Ruggenenti P, Perna A, Mosconi L, Pisoni R, Remuzzi G. Urinary protein excretion rate is the best independent predictor of ESRF in non-diabetic proteinuric chronic nephropathies. "Gruppo Italiano di Studi Epidemiologici in Nefrologia" (GISEN). *Kidney Int* 1998; **53:** 1209–16.
9 Heerspink HJ, Ninomiya T, Persson F, et al. Is a reduction in albuminuria associated with renal and cardiovascular protection? A post hoc analysis of the ALTITUDE trial. *Diabetes Obes Metab* 2016; **18:** 169–77.
10 de Zeeuw D, Remuzzi G, Parving HH, et al. Proteinuria, a target for renoprotection in patients with type 2 diabetic nephropathy: lessons from RENAAL. *Kidney Int* 2004; **65:** 2309–20.
11 Lea J, Greene T, Hebert L, et al. The relationship between magnitude of proteinuria reduction and risk of end-stage renal disease: results of the African American Study of Kidney Disease and Hypertension. *Arch Intern Med* 2005; **165:** 947–53.
12 Atkins RC, Briganti EM, Lewis JB, et al. Proteinuria reduction and progression to renal failure in patients with type 2 diabetes mellitus and overt nephropathy. *Am J Kidney Dis* 2005; **45:** 281–87.
13 Inker LA, Mondal H, Greene T, et al. Early change in urine protein as a surrogate end point in studies of IgA nephropathy: an individual-patient meta-analysis. *Am J Kidney Dis* 2016; **68:** 392–401.
14 Heerspink HJ, Kröpelin TF, Hoekman J, de Zeeuw D. Drug-induced reduction in albuminuria is associated with subsequent renoprotection: a meta-analysis. *J Am Soc Nephrol* 2015; **26:** 2055–64.
15 National Kidney Foundation. NKF-FDA-EMA Workshop. https://www.kidney.org/CKDEndpoints (accessed on Dec 31, 2018).
16 Coresh J, Heerspink HJL, Sang Y, 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* 2019; published online Jan 8. http://dx.doi.org/10.1016/S2213-8587(18)30313-9.
17 Higgins J, Green S. Cochrane handbook for systematic reviews of interventions, version 5.1.0. Cochrane Collaboration, 2011.
18 Kidney Disease: Improving Global Outcomes (KDIGO). KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. *Kidney Int Suppl* 2013; **3:** 1–150.
19 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* 2015; **175:** 1389–98.
20 Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med* 2009; **150:** 604–12.
21 Zinman B, Wanner C, Lachin JM, et al. Empagliflozin, cardiovascular outcomes, and mortality in type 2 diabetes. *N Engl J Med* 2015; **373:** 2117–28.

22 Hou FF, Zhang X, Zhang GH, et al. Efficacy and safety of benazepril for advanced chronic renal insufficiency. *N Engl J Med* 2006; **354:** 131–40.

23 Inker LA, Levey AS, Pandya K, Stoycheff N, Okparavero A, Greene T, on behalf of the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI). Early change in proteinuria as a surrogate end point for kidney disease progression: an individual patient meta-analysis. *Am J Kidney Dis* 2014; **64:** 74–85.

24 Palmer SC, Ruospo M, Teixeira-Pinto A, Craig JC, Macaskill P, Strippoli GFM. The validity of drug effects on proteinuria, albuminuria, serum creatinine, and estimated GFR as surrogate end points for ESKD: a systematic review. *Am J Kidney Dis* 2018; **72:** 779–89.

25 Inker LA, Heerspink HJ. Evaluation of surrogate end points for progression to ESKD: necessary and challenging. *Am J Kidney Dis* 2018; **72:** 771–73.

26 Blood Pressure Lowering Treatment Trialists' Collaboration. Blood pressure-dependent and independent effects of agents that inhibit the renin-angiotensin system. *J Hypertens* 2007; **25:** 951–58.

27 Cholesterol Treatment Trialists' (CTT) Collaborators. Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90 056 participants in 14 randomised trials of statins. *Lancet* 2005; **366:** 1267–78.

28 Lassere MN, Johnson KR, Schiff M, Rees D. Is blood pressure reduction a valid surrogate endpoint for stroke prevention? An analysis incorporating a systematic review of randomised controlled trials, a by-trial weighted errors-in-variables regression, the surrogate threshold effect (STE) and the Biomarker-Surrogacy (BioSurrogate) Evaluation Schema (BSES). *BMC Med Res Methodol* 2012; **12:** 27.

29 Kröpelin TF, de Zeeuw D, Andress DL, et al. Number and frequency of albuminuria measurements in clinical trials in diabetic nephropathy. *Clin J Am Soc Nephrol* 2015; **10:** 410–16.

30 Kidney Health Initiative. KHI current project: identifying surrogate endpoints for clinical trials in IgA nephropathy. American Society of Nephrology, 2017. https://www.asn-online.org/khi/project.aspx?ID=58 (accessed June 15, 2018).

31 Thompson A, Cattran DC, Blank M, Nachman PH. Complete and partial remission as surrogate end points in membranous nephropathy. *J Am Soc Nephrol* 2015; **26:** 2930–37.