Published in final edited form as:
Am J Kidney Dis. 2014 July ; 64(1): 74–85. doi:10.1053/j.ajkd.2014.02.020.

Early Change in Proteinuria as a Surrogate Endpoint for Kidney Disease Progression: An Individual Patient Meta-analysis

Lesley A. Inker, MD[*,1], **Andrew S. Levey, MD**[1], **Kruti Pandya, MS**[1], **Nicholas Stoycheff, MD**[1], **Aghogho Okparavero, MBBS**, and **Tom Greene, PhD**[2] **on behalf of the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI)**[*]

[1]Division of Nephrology, Tufts Medical Center, 800 Washington Street, Box #391, Boston, MA 02111

[2]Department of Medicine, University of Utah School of Medicine, 85 North Medical Drive East, Salt Lake City, UT 84112, USA

Abstract

Background—It is controversial whether proteinuria is a valid surrogate endpoint for randomized trials in chronic kidney disease.

Study Design—Meta-analysis of individual patient level data.

Setting & Population—Individual patient data on 9008 patients from 32 randomized trials evaluating five intervention types.

Selection Criteria for Studies—Randomized controlled trials of kidney disease progression until 2007 with measurements of proteinuria both at baseline and during the first year of follow-up, with at least one further year of follow-up for the clinical outcome.

Predictor—Early change in proteinuria.

Corresponding author and requests for single reprints: Lesley A. Inker, MD, MS, Division of Nephrology, Tufts Medical Center, 800 Washington Street, Box #391, Boston, MA 02111; Tel: +1617-636-2569; Fax: +1617-636-8329; LInker@tuftsmedicalcenter.org. A list of the CKD-EPI collaborators appears in the Acknowledgements.

Because an author of this manuscript is an editor for AJKD, the peer-review and decision-making processes were handled entirely by an Associate Editor (Jonathan C. Craig, MD, PhD, MM [Clin Epi]) who served as Acting Editor-in-Chief. Details of the journal's procedures for potential editor conflicts are given in the Information for Authors & Editorial Policies.

Financial Disclosure: Drs Inker, Levey, and Greene have an investigator-initiated grant from Pharmalink. Drs Inker and Levey have an investigator-initiated grant from Gilead Sciences. Dr Inker has consulting agreements with Otsuka. Dr Levey has an investigator initiated grant from Amgen. Dr Greene has consulting agreements with Jansen Pharmaceuticals, Keryx Biopharmaceuticals, and Genkyotech. The other authors declare that they have no other relevant financial interests.

Contributions: Study concept and design: LAI, ASL, TG; data acquisition: LAI, ASL, NS, AO, TG; data analysis and interpretation: LAI, ASL, TG, and NS; statistical analysis: LAI, ASL, TG, KP; administrative, technical, or material support: LAI, ASL, KP, NS, AO, TG; supervision: LAI, ASL, TG. Each author contributed important intellectual content during manuscript drafting or revision and accepts accountability for the overall work by ensuring that questions pertaining to the accuracy or integrity of any portion of the work are appropriately investigated and resolved. LAI takes responsibility that this study has been reported honestly, accurately, and transparently; that no important aspects of the study have been omitted, and that any discrepancies from the study as planned have been explained.

Outcomes—Doubling of serum creatinine, end stage renal disease or death.

Results—Early decline in proteinuria was associated with a lower risk of the clinical outcome (pooled HR, 0.74 per 50% reduction in proteinuria); this association was stronger at higher levels of baseline proteinuria. Pooled estimates for the proportion of treatment effect on the clinical outcome explained by early decline in proteinuria ranged from −7.0% (95% CI, −40.6% to 26.7%) to 43.9% (95% CI, 25.3% to 62.6%) across five intervention types. The direction of the pooled treatment effects on early change in proteinuria agreed with the direction of the treatment effect on the clinical outcome for all 5 intervention types, with the magnitudes of the pooled treatment effects on the two endpoints agreeing for 4 of the 5 intervention types. The pooled treatment effects on both endpoints were simultaneously stronger at higher levels of proteinuria. However, statistical power was insufficient to determine if differences in treatment effects on the clinical outcome corresponded to differences in treatment effects on proteinuria between individual studies.

Limitations—Limited variety of interventions tested and low statistical power for many chronic kidney disease clinical trials.

Conclusions—These results provide new evidence supporting the use of an early reduction in proteinuria as a surrogate endpoint, but do not provide sufficient evidence to establish its validity in all settings.

Index words

proteinuria; surrogate endpoint; kidney disease progression; disease trajectory; end-stage renal disease (ESRD); prognostic marker

---

Chronic kidney failure is a major public health issue worldwide because of its rising prevalence, poor outcomes and high cost of treatment.[1] Based on the idea that treatments initiated early in the course of a disease might slow progression and postpone the onset of kidney failure, guidelines and public health campaigns have concentrated on early detection and treatment of chronic kidney disease.[1, 2] Because many kidney diseases progress gradually, a large decline in glomerular filtration rate (GFR), assessed as a doubling of serum creatinine from baseline, is often used as a surrogate endpoint for kidney failure in randomized clinical trials (RCTs). However, the time required to reach this endpoint for patients enrolled early in the course of kidney disease often exceeds 10 years. Hence RCTs using doubling of serum creatinine as an endpoint require long durations of follow-up to detect the endpoint, increasing expense and complexity, and are infeasible for early stage disease. This problem has likely contributed to the small number of RCTs in nephrology compared to other fields, and the paucity of therapies to slow kidney disease progression.[3, 4]

The hypothesis that an early change in proteinuria is a valid surrogate endpoint for kidney disease progression in RCTs has a fairly firm biological basis.[5, 6] Proteinuria has been established as a marker of kidney damage in experimental studies and has been widely reported to be prognostic for long-term disease progression at all stages of kidney disease.[7–15] However, as evidenced by high profile past failures in other disciplines, premature acceptance of surrogate endpoints carries a risk that ineffective or harmful therapies could be approved for use in practice.[16] The National Institutes of Health and the

US Food and Drug Administration have organized several conferences to address this controversy, which had concluded that there is only preliminary empirical evidence in support of this hypothesis.[15, 17]

Here we report an individual patient-level meta-analysis of a pooled dataset of 9008 individuals from 32 RCTs to provide an integrated, systematic evaluation of an early change in proteinuria as a surrogate endpoint for trials of kidney disease progression.

Methods

A complete description of methods is included in Item S1 (provided as online supplementary materials).

Data Sources, Searches, and Study Selection

We previously described the creation of the pooled individual-level patient-level dataset.[18] In brief, we performed a systemic review of the literature for RCTs of kidney disease progression as of May 15, 2007 and requested individual patient data from the investigators. Inclusion criteria were availability of urine protein measurements at baseline and at least once within 13 months after randomization and at least one participant with a clinical outcome during one further year of follow-up. A total of 32 studies accounting for 9008 individuals which investigated five intervention types were used in the analyses reported here (A, renin angiotensin system [RAS] blockade vs. control[19–32]; B, RAS blockade vs. calcium channel blocker [CCB][19, 32–34]; C, intensive blood pressure control[19, 33, 35, 36]; D, low protein diet[35]; and E, immunosuppressive therapy[37–50]; see Table S1 for list of studies). For studies that evaluated more than one intervention,[19, 32, 33, 35] we included a separate group for each independent treatment comparison, such that some participants were included more than once. We combined the smaller studies which tested immunosuppressive therapies by disease type (IgA nephropathy, lupus nephritis and membranous nephropathy), into three separate study groups (for study-specific details, see Table S2).[37–50] Overall, we had 29 analytical comparisons (herein referred to as “studies”) across the five intervention types. We defined the active treatment as the treatment hypothesized to produce the greater reduction in the risk of the clinical endpoint.

Proteinuria

We defined an early change in proteinuria as the change in log-transformed 24-hour urine protein excretion from baseline to the first follow-up measurement between 2.5 and 13 months thereafter. We selected this interval as treatment effects on urine protein are expected to peak at about 2–4 months and some clinical trials obtained measurements only yearly. For two studies that measured urine albumin[30, 31], urine total protein was estimated from urine albumin.

Clinical Outcome

We defined the primary clinical outcome as time to the first doubling of serum creatinine, end-stage renal disease (ESRD, defined as the initiation of dialysis or transplantation), or death. We considered the composite of time to first doubling of serum creatinine or ESRD

(censoring death) in sensitivity analyses. We used the study-defined censoring times[19–25, 27, 28, 30–36, 38–41, 43–45, 48, 49], or approximated this as time from randomization to final visit date plus 6 months plus the study-specific 90th percentile of the average interval between serum creatinine measurements.[26, 29, 37, 42, 45–47, 50]

Data Synthesis and Analysis

Overview—We performed three standard categories of analyses which are widely used for validation of surrogate endpoints: 1) Association between the clinical outcome and early change in proteinuria at the individual level[51], 2) Proportion of treatment effect on the clinical outcome explained by the early change in proteinuria (Prentice-Freedman criterion)[52, 53], and 3) Association between treatment effects on the clinical outcomes and treatment effects on early change in proteinuria across different trials and/or across subgroups within trials[54–57]. For all three categories, we first obtained appropriate measures of association within each study, followed by joint analyses which summarized the results across studies. We used Bayesian mixed models for the analyses of individual level and trial level association to account for variation between trials when summarizing overall results[55, 58, 59]. We used credible intervals, which are in some respects analogous to confidence intervals in frequentist statistics, to characterize the precision of parameter estimates from Bayesian analyses[59].

Individual-Level Association—Demonstration of a consistent patient-level epidemiologic association between a surrogate and the clinical outcome is widely regarded as necessary, although not sufficient, for establishing the validity of the surrogate endpoint in clinical trials[60–62]. We evaluated individual-level association by performing separate Cox regressions to relate the clinical outcome to early change in proteinuria in each study, with results expressed as the hazard ratio associated with a halving of proteinuria. The primary analyses were adjusted for baseline proteinuria. Additional models adjusted for age, sex, baseline serum creatinine and mean arterial pressure in addition to proteinuria. The study-specific results were subsequently analyzed under Bayesian mixed effect models to summarize the distribution of individual level association across all studies, within each of the five interventions, and in relation to the level of baseline proteinuria[58]. For the pooled result across all studies, we only included one intervention per study, such that participants were not represented more than once.

Proportion of Treatment Effect Explained (Prentice-Freedman Criterion)—The proportion of the treatment effect on a clinical outcome “explained by the surrogate” has been widely used as an index of the validity of surrogate endpoints[52, 53, 63]. The proportion of treatment effect is defined as the ratio of the treatment effect on the clinical outcome that remains after statistically controlling for the surrogate to the treatment effect without controlling for the surrogate. Large proportions of treatment effect close to 1 are regarded as supporting the surrogacy hypothesis[54, 64].

We performed Cox regressions to estimate the treatment effects on the clinical outcomes for each study, first adjusting only for baseline proteinuria. Then for studies in which the treatment effect on the clinical outcome approached statistical significance (p-value< 0.10),

we repeated the Cox regression adjusting also for the early change in proteinuria. The proportion of treatment effect was calculated as 1 minus the ratio of the log-transformed Cox regression coefficients for the treatment with and without adjusting for early change in proteinuria. These analyses were repeated with additional adjustment for the extended covariate set described above.

Trial-Level Analyses—Assessments of individual-level association and the Prentice-Freedman criteria both depend on the untestable assumption of no residual confounding from factors which jointly influence the surrogate and clinical endpoints[54, 64]. By contrast, trial-level analyses investigate the relationship of treatment effects on the surrogate with treatment effects on the clinical endpoints, where each treatment effect is estimated from a randomized comparison, and therefore minimizes the risk of confounding that affects the first two approaches[54]. Trial-level analyses require heterogeneity to be informative; demonstration of treatment effects on the surrogate that agree with those of the clinical endpoint across studies across varying treatment effects on the surrogate and a wide range of interventions increases confidence that the treatment effect on the surrogate will predict the treatment effect on the clinical outcome in future RCTs, supporting the surrogacy hypothesis.

The first step for all the trial-level analyses was to apply linear and Cox regression separately in each study to estimate the treatment effects on early change in proteinuria (expressed as the ratio of follow-up vs. baseline geometric mean proteinuria between treatment groups) and on the clinical outcome (expressed as hazard ratios [HRs]), and to estimate terms characterizing the interactions of these treatment effects with baseline proteinuria. We sought to capitalize on heterogeneity by analyzing these quantities in three different ways.

We first considered *variation across the five interventions* by applying Bayesian mixed models to obtain pooled estimates of the treatment effects on each endpoint for each intervention type, and then computing ratios, or relative effects, between the pooled estimates for the two endpoints[51]. Consistent ratios across the five interventions would suggest an agreement of the treatment effect on the surrogate and on the clinical outcome that is independent of the mechanism used to lower proteinuria, supporting the surrogacy hypothesis.

Our second approach focused on *variation in treatment effects on early change in proteinuria among studies* by applying a Bayesian mixed effect regression model to relate treatment effects on the clinical outcome to treatment effects on early change in proteinuria, with study as the unit of analysis. A regression slope substantially greater than zero would indicate that larger treatment effects on early change in proteinuria accurately predict larger treatment effects on the clinical endpoint and support the surrogacy hypothesis. This approach has been a primary focus of the statistical surrogate endpoint literature[55–57, 65, 66].

Our third approach sought to capitalize on *variation in treatment effects across different levels of baseline proteinuria* which has been reported for several interventions[19, 67–69]. Bayesian mixed effects regression analyses including interaction terms with baseline

proteinuria were used to assess if the treatment effects on the two endpoints varied in similar way between different baseline proteinuria levels. These analyses were repeated with proteinuria defined as a continuous and as a categorical variable (baseline proteinuria ≤1, >1–≤3, and > 3 g/d), and with and without adjustment for study and intervention type.

Variation of Results Among Studies—We summarized the variation among studies of individual-level association HRs and of the geometric mean ratios and HRs for treatment effects on change in proteinuria and the clinical outcome by reporting the projected range of the middle 90% of HRs or geometric mean ratios across studies which are implied by the posterior median standard deviation under the Bayesian models.[70] Formal assessments of the evidence that variation exceeds 0 are based on 95% Bayesian credible intervals for the standard deviations of the log-transformed HRs and geometric mean ratios. Variation in HRs and geometric mean ratios among interventions is assessed by the 95% Bayesian credible intervals for comparisons between each intervention and the RAS blockade vs. control intervention.

Results

Dataset and Composite Events

Table 1 and Table S2–Table S3 show characteristics of the studies and patients. The dataset included 9008 people, across five interventions types: RAS blockade vs. control (5748 people), RAS blockade vs. CCB (2295 people), intensive blood pressure lowering (2655 people), low protein diet (839 people), and immunosuppressive therapy (804 people). Over a median of 2.65 years' follow-up, there were 5146 composite events (2031 doublings of serum creatinine, 1981 cases of ESRD, and 1134 deaths). Reflecting the predominance of events related to kidney disease progression (78% of all events) in the composite, sensitivity analyses censoring death produced similar results to analyses reported below based on the full composite outcome.

Individual Level Association

Figure 1 and Table S4 shows the association of the early change in proteinuria with the subsequent clinical endpoint. An early decline in proteinuria was consistently associated with a lower risk for the clinical outcome (pooled HR per 50% reduction in proteinuria, 0.74; 95% Bayesian credible interval, 0.67–0.82), although the magnitude of the association varied moderately between studies (HR, 0.56–0.98 for 90% of studies) (Table S5). Similar results were obtained with extended covariate adjustment (Figure 1 and Table S4). The individual level association was significantly stronger at higher levels of baseline proteinuria, with halving of proteinuria leading to a 12% reduced hazard of the clinical outcome for baseline proteinuria <1 g/d vs. a 31% reduction for baseline proteinuria > 3 g/d (Table 2). This dependence of individual level association on baseline proteinuria persisted after adjustment for study, and hence treatment type (Table S6).

Proportion of Treatment Effect Explained (Prentice- Freedman criteria)

Figure 2 and Table S7 show the treatment effects on the clinical endpoint before and after adjusting for proteinuria and the associated proportion of treatment effect. The pooled

proportions of treatment effect for the five intervention types range from −7.0% (95% confidence interval, −40.6% to 26.7%) for studies of immunosuppressive therapy (indicating slightly larger treatment effects *after* adjustment for early change in proteinuria) to 43.9% (95% confidence interval, 25.3% to 62.6%) for studies of RAS blockade vs. placebo (indicating smaller treatment effects *after* adjustment for early change in proteinuria).

Trial Level Analyses

Figure 3 and Table S8 show the treatment effects on early change in proteinuria and the clinical outcome. Treatment was generally associated with a greater reduction in proteinuria compared to control (pooled geometric mean ratio, 0.77; 95% Bayesian credible interval, 0.72–0.82), with moderately large variation among studies (geometric mean ratio, 0.60–0.98 across 90% of studies) (Figure 3 left panel and Table S5 and Table S8). Treatment led to an improved clinical outcome compared to control across all studies (pooled HR, 0.79; 95% credible interval, 0.73 – 0.86)], with little variation among studies (HRs, 0.74 to 0.85 across 90% of studies) (Figure 3 [right panel] and Table S5 and Table S8). Bayesian credible intervals indicated no clear evidence of differences in pooled treatment effects among the five intervention types on either early change in proteinuria or on the clinical outcome (Figure 3 [footnote]).

As shown in Figure 4 (top panel), the directions of the pooled treatment effects on early change in proteinuria and on the clinical outcome agreed for all five treatment comparison classes, with each treatment reducing both proteinuria and the risk of the clinical composite, although not significantly in all cases. The ratios of the pooled treatment effects across intervention types were consistent with each other (range, 0.95 (95% Bayesian credible interval, 0.68–1.98) to 1.08 (95% Bayesian credible interval, 0.92–2.50)], except for immunosuppressive therapy vs. control, in which the lower ratio (0.72 [95% Bayesian credible interval, 0.46–1.58)] was largely the result of a beneficial pooled treatment effect on the clinical outcome but not early change in proteinuria in some of the IgA nephropathy studies (Table S8).

Figure 4 (bottom panel) shows the relationship between treatment effects on the clinical outcome vs. the treatment effects on early change in proteinuria across individual studies. The slope of the regression line relating the treatment effects on the two endpoints was not estimated with sufficient precision to be informative. The 95% Bayesian credible interval ranges from 0.60% lower to 0.64% higher HR for the clinical endpoint associated with a 1% lower geometric mean ratio for proteinuria. This indicates that there was insufficient variation in treatment effects on change in proteinuria and/or insufficient statistical power in the bulk of the studies to determine whether differences in treatment effects on proteinuria are associated with similar differences in treatment effects on the clinical outcome. In sensitivity analyses, similar results were seen when treatment effects on proteinuria were evaluated in absolute units of g/d.

Table 2 shows the variation in the treatment effects on the two endpoints in relation to baseline proteinuria. When change in proteinuria was modeled as a relative change, in patients with proteinuria < 1 g/d, treatment led to an average 25% reduction in proteinuria but no discernible reduction in clinical events. However, when change in proteinuria was

modeled in units of g/d, average treatment effects on both the clinical outcome and change in proteinuria were negligible when baseline proteinuria was < 1 g/d but increased substantially at higher levels of baseline proteinuria. This concordant pattern of larger treatment effects on both endpoints at higher levels of baseline proteinuria but smaller effects on both endpoints at lower baseline proteinuria persisted after adjustment for study, and hence treatment type (Table S6).

Discussion

Use of valid surrogate endpoints may improve the efficiency of clinical trials. However, not all surrogates are valid and there are numerous examples of discrepancies between treatment effects on the surrogate and clinical endpoint.[16] Proteinuria could be a useful surrogate for trials of CKD progression since it is often occurs early in the course of disease and can be measured frequently and inexpensively. Findings from experimental studies indicate an important function for proteinuria in the pathogenesis of kidney disease progression, suggesting that an early change in proteinuria may be a valid surrogate endpoint for clinical trials of interventions hypothesized to reduce proteinuria.[7, 31] This report provides a comprehensive evaluation of this hypothesis based on a joint analysis of over 9000 individuals from 32 RCTs of five types of interventions in progressive kidney disease. Strengths of this study include a systematic literature search to include all available studies until 2007, uniform definitions of exposures and outcomes, and a comprehensive evaluation using the three standard approaches for validating surrogate endpoints in the statistical and medical literatures. The results from these analyses extend the evidence supporting use of proteinuria in some settings.

Our analyses of individual level association established that greater early reduction in proteinuria is consistently associated with slower progression of kidney disease across all five interventions and this association was stronger when baseline proteinuria was higher, although it varied moderately among studies. These results are limited by possible confounding by factors that influence both the surrogate and the clinical endpoint but the results were little changed after adjustment for a limited set of baseline covariates. Our results are consistent with and extend results of epidemiologic studies and observational analyses of clinical trials that demonstrated the utility of proteinuria as a prognostic marker for subsequent clinical outcomes, and they support the use of change in proteinuria to inform prognosis in clinical practice[7–15, 18]

The proportion of treatment effect is a traditional method to evaluate surrogate endpoints but subject to bias due to measurement in error in proteinuria and as well as possible residual confounding.[54, 64] Our assessments of the Prentice-Freedman criteria were inconclusive, with proportion of treatment effect ranging from slightly negative in studies of immunosuppressive therapy to moderately positive for comparisons of RAS blockade, with wide confidence intervals for all interventions. As in the individual level analyses, the results were little changed after adjustment for covariates, but the risk of confounding remains. Our interpretation is that these analyses do not provide support either for or against proteinuria as a surrogate endpoint.

We used three trial-level approaches to investigate if treatment effects on change in proteinuria agreed with treatment effects on the clinical outcome. In the first approach, we found that pooled estimates of treatment effects were consistent with reductions both in proteinuria and in the risk of the clinical outcome for each of the five intervention types. Similar analyses using less formal methods have been interpreted as supporting reduction in blood pressure and serum cholesterol as surrogate endpoints for cardiovascular disease protection.[71–73] In the third approach, we showed that the treatment effects on both proteinuria and the clinical outcome were significantly greater for higher vs. lower levels of baseline proteinuria. This finding is consistent with experimental studies showing greater effects of therapies to lower proteinuria in proteinuric kidney diseases and with the hypothesis that treatment effects on change in proteinuria are predictive of treatment effects on kidney disease progression.[7, 31] Both of these trial level results support the surrogacy hypothesis. The second approach, which related the size of treatment effects on the two endpoints across different trials, was uninformative as the regression slope relating the treatment effects was non-significant but with a confidence interval too wide to rule out a strong relationship. Many of the available studies were small, and the larger studies were mostly of studies of a single treatment type, RAS blockade. Hence, there is not sufficient variation in treatment effects among well powered CKD trials to determine whether or not different estimated treatment effects on early change in proteinuria are predictive of different treatment effects on the clinical outcome.

Our analysis has some limitations. First, our designation of the treatment arm in each trial as the group hypothesized to have the greater benefit was somewhat arbitrary. Second, limitations in sample size of past clinical trials and in variation among trials in treatment effects on proteinuria limited statistical power, particularly for trial-level analyses. Thus, the evidence suggesting agreement of treatment effects on the two endpoints within the five interventions is limited by the imprecision in the pooled estimated treatment effects for several of interventions, and the estimated effects did not differ significantly from 0 in several cases. The limited number of large trials also means that our estimates of variation in treatment effects were influenced by our assumptions for the prior distributions for variation in parameters among studies. However, we included all clinical trials that met our pre-specified eligibility criteria and we carefully developed the priors based on clinical knowledge. Third, inclusion of death as a component of the clinical composite outcome introduces non-kidney events in the analyses. However, we found similar results in sensitivity analyses excluding death from the composite. Fourth, our analyses are restricted to the specific diseases and interventions included in the published and unpublished that were included at the beginning of our study in 2007. Inclusion of additional trials, in particular, large trials of new interventions and interventions that have negative results, could overcome some of the limitations of our current analysis. However, interventions that do not have positive effects on proteinuria in early phase clinical trials are frequently not evaluated in Phase III clinical trials, and some Phase III trials without a beneficial effect on proteinuria have been terminated prior to completion, precluding the ability to relate the treatment effect on proteinuria to the treatment effect on the clinical outcome[74]. Fifth, our evaluation of proteinuria as a surrogate endpoint was limited to changes between 2.5 and 13 months, and may not apply to surrogate endpoints defined by changes in proteinuria over

longer periods. Finally, the estimates of heterogeneity between trials reflects not only biological variation in treatment effects, but also variation in designs and procedures between trials.

In summary, the evidence presented here is not sufficient to conclude that treatment effects on early change in proteinuria reliably predict treatment effects on clinical outcome in all circumstances. However, due to limitations of the data included, this conclusion should not be misconstrued as a refutation of the potential validity of proteinuria as a surrogate endpoint for kidney disease progression. Indeed, when considered in conjunction with evidence from experimental studies, we believe the findings from our analyses are sufficient to recommend continued use of proteinuria as a surrogate endpoint in early-phase clinical trials for new therapies and for exploratory analyses (e.g., subgroup analyses). In addition, in kidney diseases and populations where proteinuria is high and experimental evidence for a pathological role of proteinuria is particularly strong, cautious use of proteinuria as a surrogate endpoint may be warranted in certain Phase III clinical trials, especially when the risks of adverse outcomes of the intervention are low and there is no alternative (e.g. rare or infrequent diseases where adequate sample sizes for clinical outcomes are infeasible). For populations with high levels of proteinuria and high GFR, it may also be reasonable to use reduction in proteinuria for initial acceptance of an intervention, with subsequent postapproval conformation of the treatment effect on the clinical outcome. Further delineation of the scope for appropriate use of proteinuria as a surrogate endpoint in clinical trials would require additional data from large well powered trials across a broad array of treatment classes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The CKD-EPI collaborators contributed data and reviewed the manuscript, and are as follows: African American Study of Kidney Disease and Hypertension (AASK),Tom Greene, PhD, Utah University, Salt Lake City, Utah; ACE Inhibition in Progressive Renal Disease (AIPRD) Study, Robert Toto, MD, University of Texas Southwestern Medical Center; GG van Essen, MD, University Hospital Groningen, Netherlands; the Angiotensin-Converting-Enzyme Inhibition in Progressive Renal Insufficiency Study Group, Giuseppe Maschio, MD, University Hospital, Verona, Italy; Anne Lise Kamper, Svend Strandgaard, University of Copenhagen, Denmark; Barry M. Brenner, Brigham and Women's Hospital, Boston, Massachusetts; Benno U. Ihle MBBS, FRACP, FJFICM, Priscilla S. Kincaid-Smith MD, DSc, Royal Melbourne Hospital, Australia; Thierry P. Hannedouche, MD, University Hospital, Strasbourg, France; Hong Kong Study Using Valsartan in IgA Nephropathy (HKVIN), Philip Kam-Tao Li, MD, FRCP, FACP, Prince of Wales Hospital, Chinese University of Hong Kong, Shatin, Hong Kong; Fan FanHou, MD, PhD, Nanfang Hospital, Southern Medical University, Guangzhou, China; Manuel Praga, MD, Hospital 12 de Octubre, Madrid, Spain; Captopril in Diabetic Nephropathy Study (CSG), Roger A. Rodby, MD, Richard D. Rohde, BS, Rush-Presbyterian-St. Luke's Medical Center, Chicago, Illinois; Reduction of Endpoints in NIDDM with the Angiotensin II Antagonist Losartan (RENAAL), Dick De Zeeuw, MD, PhD, H.J. Lambers Heerspink PharmD PhD, Barry M. Brenner, MD, William Keane, MD, Brigham and Women's Hospital, Boston, Massachusetts; Ramipril Efficacy In Nephropathy (REIN), Giuseppe Remuzzi, MD, FRCP, Piero Ruggenenti MD, Mario Negri Institute for Pharmacological Research, Clinical Research Centre for Rare Diseases, Ranica, Italy; Irbesartan in Diabetic Nephropathy Trial (IDNT), Ed Lewis, MD, Lawrence G. Hunsicker, MD, Rush-Presbyterian-St. Luke's Medical Center, Chicago, Illinois; Appropriate Blood Pressure Control in Diabetes (ABCD), Robert W. Schrier, MD, Raymond O. Estacio, MD, University of Colorado Health Sciences Center, Denver; Pietro Zucchelli, MD, Ospedale M. Malpighi, Bologna, Italy; Modification of Diet in Renal Disease (MDRD) Study, Gerald Beck, PhD, Cleveland Clinic Foundation, Ohio; Ramipril Efficacy In Nephropathy (REIN-2), Giuseppe Remuzzi, MD, FRCP, Piero Ruggenenti, MD, Mario Negri Institute for Pharmacological Research, Clinical Research Centre for Rare Diseases, Ranica, Italy; Bart Maes, MD, PhD, University Hospital Gasthuisberg, Leuven, Belgium; James

Donadio, MD, Fernando Fervenza, MD, Mayo Clinic, Rochester, Minnesota; Gerald B. Appel, MD, Gershon Frisch, MD, New York Presbyterian Hospital, Columbia University, New York; Euro Lupus Nephritis Trial (ELNT), Frédéric A. Houssiau, MD, PhD, Cliniques Universitaires St-Luc and Institut de Recherche Expérimentale et Clinique, Université catholique de Louvain, Belgium; Tak-Mao Chan MD, University of Hong Kong, Queen Mary Hospital, Pokfulam, Hong Kong; Lupus Nephritis Collaborative Study (LNCS), Edmund Lewis, MD, John M. Lachin, ScD, Rush-Presbyterian-St. Luke's Medical Center, Chicago, Illinois; Daniel C. Cattran, MD, FRCPC, University of Toronto, Canada; Claudio Ponticelli, MD, Patrizia Passerini, MD, Gabriella Moroni, MD, Giuseppe Montogrino, MD, IRCCS Istituto Humanitas, Milan, Italy; Manuel Praga, MD, Grupo Español de estudio de la Nefropatía Membranosa; Tazeen H. Jafar, Duke-NUS Graduate Medical School, Singapore, Aga Khan University, Karachi, Pakistan.

Support: This study is supported by grants UO1 DK 053869 and UO1 DK 35073 from National Institutes of Health (NIH). The study was funded by a cooperative agreement with the NIH, which allows them substantial involvement in the design of the study and in the collection, analysis, and interpretation of the data. The funding source was not required to approve publication of the finished manuscript.

References

1. Levey AS, Coresh J. Chronic kidney disease. Lancet. 2012; 379(9811):165–180. [PubMed: 21840587]
2. National Kidney Foundation. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. Am J Kidney Dis. 2002; 39(2 Suppl 1):S1–S266. [PubMed: 11904577]
3. Strippoli GF, Craig JC, Schena FP. The number, quality, and coverage of randomized controlled trials in nephrology. J Am Soc Nephrol. 2004; 15(2):411–419. [PubMed: 14747388]
4. Palmer SC, Sciancalepore M, Strippoli GF. Trial quality in nephrology: how are we measuring up? Am J Kidney Dis. 2011; 58(3):335–337. [PubMed: 21856491]
5. Burton C, Harris KP. The role of proteinuria in the progression of chronic renal failure. Am J Kidney Dis. 1996; 27(6):765–775. [PubMed: 8651239]
6. Remuzzi G, Chiurchiu C, Ruggenenti P. Proteinuria predicting outcome in renal disease: nondiabetic nephropathies (REIN). Kidney Int Suppl. 2004; 92:S90–S96. [PubMed: 15485427]
7. Remuzzi G, Benigni A, Remuzzi A. Mechanisms of progression and regression of renal lesions of chronic nephropathies and diabetes. J Clin Invest. 2006; 116(2):288–296. [PubMed: 16453013]
8. de Zeeuw D, Remuzzi G, Parving H-H, et al. Proteinuria, a target for renoprotection in patients with type 2 diabetic nephropathy: Lessons from RENAAL. Kidney Int. 2004; 65(6):2309–2320. [PubMed: 15149345]
9. Hunsicker LG, Adler S, Caggiula A, et al. Predictors of the progression of renal disease in the Modification of Diet in Renal Disease Study. Kidney Int. 1997; 51(6):1908–1919. [PubMed: 9186882]
10. 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(8):947–953. [PubMed: 15851648]
11. Rossing K, Christensen PK, Hovind P, Tarnow L, Rossing P, Parving HH. Progression of nephropathy in type 2 diabetic patients. Kidney Int. 2004; 66(4):1596–1605. [PubMed: 15458456]
12. 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(5):1209–1216. [PubMed: 9573535]
13. Hunsicker LG, Atkins RC, Lewis JB, et al. Impact of irbesartan, blood pressure control, and proteinuria on renal outcomes in the Irbesartan Diabetic Nephropathy Trial. Kidney Int Suppl. 2004; 92:S99–S101. [PubMed: 15485429]
14. Jafar TH, Stark PC, Schmid CH, et al. Proteinuria as a modifiable risk factor for the progression of non-diabetic renal disease. Kidney Int. 2001; 60(3):1131–1140. [PubMed: 11532109]
15. Levey AS, Cattran D, Friedman A, et al. Proteinuria as a surrogate outcome in CKD: report of a scientific workshop sponsored by the National Kidney Foundation and the US Food and Drug Administration. Am J Kidney Dis. 2009; 54(2):205–226. [PubMed: 19577347]

16. Echt D, Liebson R, Mitchell N. Mortality and morbidity in patient receiving encainide, flecainide, or placebo. The Cardiac Arrhythmia Suppression Trial. N Engl J Med. 1991; 324:781–788. [PubMed: 1900101]
17. Stevens LA, Greene T, Levey AS. Surrogate end points for clinical trials of kidney disease progression. Clin J Am Soc Nephrol. 2006; 1(4):874–884. [PubMed: 17699300]
18. Stoycheff N, Pandya K, Okparavero A, et al. Early change in proteinuria as a surrogate outcome in kidney disease progression: a systematic review of previous analyses and creation of a patient-level pooled dataset. Nephrol Dial Transplant. 2011; 26(3):848–857. [PubMed: 20817671]
19. Wright JT Jr. Bakris G, Greene T, et al. Effect of blood pressure lowering and antihypertensive drug class on progression of hypertensive kidney disease: results from the AASK trial. JAMA. 2002; 288(19):2421–2431. [PubMed: 12435255]
20. Jafar TH, Stark PC, Schmid CH, et al. Progression of chronic kidney disease: the role of blood pressure control, proteinuria, and angiotensin-converting enzyme inhibition: a patient-level meta-analysis. Ann Intern Med. 2003; 139(4):244–252. [PubMed: 12965979]
21. van Essen GG, Apperloo AJ, Rensma PL, et al. Are angiotensin converting enzyme inhibitors superior to beta blockers in retarding progressive renal function decline? Kidney Int Suppl. 1997; 63:S58–S62. [PubMed: 9407423]
22. Maschio G, Alberti D, Janin G, et al. Effect of the angiotensin-converting-enzyme inhibitor benazepril on the progression of chronic renal insufficiency. The Angiotensin-Converting-Enzyme Inhibition in Progressive Renal Insufficiency Study Group. N Engl J Med. 1996; 334(15):939–945. [PubMed: 8596594]
23. Kamper AL, Strandgaard S, Leyssac PP. Effect of enalapril on the progression of chronic renal failure. A randomized controlled trial. Am J Hypertens. 1992; 5(7):423–430. [PubMed: 1637513]
24. Ihle BU, Whitworth JA, Shahinfar S, Cnaan A, Kincaid-Smith PS, Becker GJ. Angiotensin-converting enzyme inhibition in nondiabetic progressive renal insufficiency: a controlled double-blind trial. Am J Kidney Dis. 1996; 27(4):489–495. [PubMed: 8678058]
25. Hannedouche T, Landais P, Goldfarb B, et al. Randomised controlled trial of enalapril and beta blockers in non-diabetic chronic renal failure. BMJ. 1994; 309(6958):833–837. [PubMed: 7950612]
26. Li PK, Leung CB, Chow KM, et al. Hong Kong study using valsartan in IgA nephropathy (HKVIN): a double-blind, randomized, placebo-controlled study. Am J Kidney Dis. 2006; 47(5): 751–760. [PubMed: 16632013]
27. Hou FF, Zhang X, Zhang GH, et al. Efficacy and safety of benazepril for advanced chronic renal insufficiency. N Engl J Med. 2006; 354(2):131–140. [PubMed: 16407508]
28. Praga M, Gutierrez E, Gonzalez E, Morales E, Hernandez E. Treatment of IgA nephropathy with ACE inhibitors: a randomized and controlled trial. J Am Soc Nephrol. 2003; 14(6):1578–1583. [PubMed: 12761258]
29. Lewis EJ, Hunsicker LG, Bain RP, Rohde RD. The effect of angiotensin-converting-enzyme inhibition on diabetic nephropathy. The Collaborative Study Group. N Engl J Med. 1993; 329(20): 1456–1462. [PubMed: 8413456]
30. Brenner BM, Cooper ME, de Zeeuw D, et al. Effects of losartan on renal and cardiovascular outcomes in patients with type 2 diabetes and nephropathy. N Engl J Med. 2001; 345(12):861–869. [PubMed: 11565518]
31. Ruggenenti P, Perna A, Gherardi G, et al. Renoprotective properties of ACE-inhibition in non-diabetic nephropathies with non-nephrotic proteinuria. Lancet. 1999; 354(9176):359–364. [PubMed: 10437863]
32. Lewis EJ, Hunsicker LG, Clarke WR, et al. Renoprotective effect of the angiotensin-receptor antagonist irbesartan in patients with nephropathy due to type 2 diabetes. New England Journal Medicine. 2001; 345(12):851–860.
33. Estacio RO, Jeffers BW, Gifford N, Schrier RW. Effect of blood pressure control on diabetic microvascular complications in patients with hypertension and type 2 diabetes. Diabetes Care. 2000; 23(Suppl 2):B54–B64. [PubMed: 10860192]
34. Zucchelli P, Zuccala A, Borghi M, et al. Long-term comparison between captopril and nifedipine in the progression of renal insufficiency. Kidney Int. 1992; 42(2):452–458. [PubMed: 1405330]

35. Klahr S, Levey AS, Beck GJ, et al. The effects of dietary protein restriction, blood-pressure control on the progression of chronic renal disease. Modification of Diet in Renal Disease Study Group. N Engl J Med. 1994; 330(13):877–884. [PubMed: 8114857]
36. Ruggenenti P, Perna A, Loriga G, et al. Blood-pressure control for renoprotection in patients with non-diabetic chronic renal disease (REIN-2): multicentre, randomised controlled trial. Lancet. 2005; 365(9463):939–946. [PubMed: 15766995]
37. Maes BD, Oyen R, Claes K, et al. Mycophenolate mofetil in IgA nephropathy: results of a 3-year prospective placebo-controlled randomized study. Kidney Int. 2004; 65(5):1842–1849. [PubMed: 15086925]
38. Donadio JV Jr. Larson TS, Bergstralh EJ, Grande JP. A randomized trial of high-dose compared with low-dose omega-3 fatty acids in severe IgA nephropathy. J Am Soc Nephrol. 2001; 12(4): 791–799. [PubMed: 11274240]
39. Donadio JV Jr. Grande JP, Bergstralh EJ, Dart RA, Larson TS, Spencer DC. The long-term outcome of patients with IgA nephropathy treated with fish oil in a controlled trial. Mayo Nephrology Collaborative Group. J Am Soc Nephrol. 1999; 10(8):1772–1777. [PubMed: 10446945]
40. Frisch G, Lin J, Rosenstock J, et al. Mycophenolate mofetil (MMF) vs placebo in patients with moderately advanced IgA nephropathy: a double-blind randomized controlled trial. Nephrol Dial Transplant. 2005; 20(10):2139–2145. [PubMed: 16030050]
41. Houssiau FA, Vasconcelos C, D'Cruz D, et al. Immunosuppressive therapy in lupus nephritis: the Euro-Lupus Nephritis Trial, a randomized trial of low-dose versus high-dose intravenous cyclophosphamide. Arthritis and rheumatism. 2002; 46(8):2121–2131. [PubMed: 12209517]
42. Chan TM, Tse KC, Tang CS, Mok MY, Li FK. Long-term study of mycophenolate mofetil as continuous induction and maintenance treatment for diffuse proliferative lupus nephritis. J Am Soc Nephrol. 2005; 16(4):1076–1084. [PubMed: 15728784]
43. Lewis EJ, Hunsicker LG, Lan SP, Rohde RD, Lachin JM. A controlled trial of plasmapheresis therapy in severe lupus nephritis. The Lupus Nephritis Collaborative Study Group. N Engl J Med. 1992; 326(21):1373–1379. [PubMed: 1569973]
44. Cattran DC, Appel GB, Hebert LA, et al. A randomized trial of cyclosporine in patients with steroid-resistant focal segmental glomerulosclerosis. North America Nephrotic Syndrome Study Group. Kidney Int. 1999; 56(6):2220–2226. [PubMed: 10594798]
45. Cattran DC, Appel GB, Hebert LA, et al. Cyclosporine in patients with steroid-resistant membranous nephropathy: a randomized trial. Kidney Int. 2001; 59(4):1484–1490. [PubMed: 11260412]
46. Ponticelli C, Zucchelli P, Passerini P, et al. A randomized trial of methylprednisolone and chlorambucil in idiopathic membranous nephropathy. N Engl J Med. 1989; 320(1):8–13. [PubMed: 2642605]
47. Ponticelli C, Zucchelli P, Passerini P, Cesana B. Methylprednisolone plus chlorambucil as compared with methylprednisolone alone for the treatment of idiopathic membranous nephropathy. The Italian Idiopathic Membranous Nephropathy Treatment Study Group. N Engl J Med. 1992; 327(9):599–603. [PubMed: 1640953]
48. Ponticelli C, Altieri P, Scolari F, et al. A randomized study comparing methylprednisolone plus chlorambucil versus methylprednisolone plus cyclophosphamide in idiopathic membranous nephropathy. J Am Soc Nephrol. 1998; 9(3):444–450. [PubMed: 9513907]
49. Ponticelli C, Passerini P, Salvadori M, et al. A randomized pilot trial comparing methylprednisolone plus a cytotoxic agent versus synthetic adrenocorticotropic hormone in idiopathic membranous nephropathy. Am J Kidney Dis. 2006; 47(2):233–240. [PubMed: 16431252]
50. Praga M, Barrio V, Juarez GF, Luno J. Tacrolimus monotherapy in membranous nephropathy: a randomized controlled trial. Kidney Int. 2007; 71(9):924–930. [PubMed: 17377504]
51. Buyse M, Molenberghs G. Criteria for the Validation of Surrogate Endpoints in Randomized Experiments. Biometrics. 1998; 54:1014–1029. [PubMed: 9840970]
52. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. Statistics in medicine. 1989; 8(4):431–440. [PubMed: 2727467]

53. Freedman L, Graubard B, Schatzkin L. Statistical Validation of Intermediate Endpoints for Chronic Diseases. Statistics in medicine. 1992; 11:167–178. [PubMed: 1579756]
54. Joffe MM, Greene T. Related causal frameworks for surrogate outcomes. Biometrics. 2009; 65(2): 530–538. [PubMed: 18759836]
55. Daniels MJ, Hughes MD. Meta-analysis for the evaluation of potential surrogate markers. Statistics in medicine. 1997; 16(17):1965–1982. [PubMed: 9304767]
56. Gail M, Pfeiffer R, van Houwelingen H, Carroll R. On Meta-Analytic Assessment of Surrogate Outcomes. Biostatistics. 2000; 1:231–246. [PubMed: 12933506]
57. Buyse M, Molenberghs G, Burzykowski T, Renard D, Geys H. The validation of surrogate endpoints in meta-analyses of randomized experiments. Biostatistics. 2000; 1:49–68. [PubMed: 12933525]
58. Abrams, KR.; Jones, DR.; Jones, DR.; Sheldon, TA.; Song, F. Methods for Meta-Analysis in Medical Research. J. Wiley; 2000. Bayesian Methods in Meta-Analysis; p. 163-190.
59. Congdon, P. Bayesian statistical modeling. New York: Wiley; 2007. Multilevel and Panel Data Models; p. 367-424.
60. Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? Ann Intern Med. 1996; 125(7):605–613. [PubMed: 8815760]
61. Baker SG, Kramer BS. A perfect correlate does not a surrogate make. BMC medical research methodology. 2003; 3:16. [PubMed: 12962545]
62. Katz R. Biomarkers and surrogate markers: an FDA perspective. NeuroRx : the journal of the American Society for Experimental NeuroTherapeutics. 2004; 1(2):189–195. [PubMed: 15717019]
63. Lin D, Fleming T, De Gruttola V. Estimating the proportion of treatment effect explained by a surrogate marker. Statistics in medicine. 1997; 16:1515–1527. [PubMed: 9249922]
64. Frangakis C, Rubin D. Principal stratification and causal inference. Biometrics. 2002; 58:21–29. [PubMed: 11890317]
65. Burzykowski, T.; Molenberghs, G.; Buyse, M., editors. The Evaluation of Surrogate Endpoints. New York: Springer; 2005.
66. Korn EL, Albert PS, McShane LM. Assessing surrogates as trial endpoints using mixed models. Statistics in medicine. 2005; 24(2):163–182. [PubMed: 15515150]
67. Jafar TH, Schmid CH, Landa M, et al. Angiotensin-converting enzyme inhibitors progression of nondiabetic renal disease. A meta-analysis of patient-level data. Ann Intern Med. 2001; 135(2): 73–87. [PubMed: 11453706]
68. Peterson JC, Adler S, Burkart JM, et al. Blood pressure control, proteinuria, the progression of renal disease. The Modification of Diet in Renal Disease Study. Ann Intern Med. 1995; 123(10): 754–762. [PubMed: 7574193]
69. Lv J, Ehteshami P, Sarnak MJ, et al. Effects of intensive blood pressure lowering on the progression of chronic kidney disease: a systematic review and meta-analysis. CMAJ. 2013; 185(11):949–957. [PubMed: 23798459]
70. Steel PD, Kammeyer-Mueller J. Bayesian Variance Estimation for Meta-Analysis Quantifying Our Uncertainty. Organizational Research Methods. 2008; 11(1):54–78.
71. Czernichow S, Zanchetti A, Turnbull F, et al. The effects of blood pressure reduction and of different blood pressure-lowering regimens on major cardiovascular events according to baseline blood pressure: meta-analysis of randomized trials. J Hypertens. 2011; 29(1):4–16. [PubMed: 20881867]
72. Turnbull F, Neal B, Pfeffer M, et al. Blood pressure-dependent and independent effects of agents that inhibit the renin-angiotensin system. J Hypertens. 2007; 25(5):951–958. [PubMed: 17414657]
73. Baigent C, Keech A, Kearney PM, et al. Efficacy and safety of cholesterollowering treatment: prospective meta-analysis of data from 90,056 participants in 14 randomised trials of statins. Lancet. 2005; 366(9493):1267–1278. [PubMed: 16214597]
74. Lewis EJ, Lewis JB, Greene T, et al. Sulodexide for kidney protection in type 2 diabetes patients with microalbuminuria: a randomized controlled trial. Am J Kidney Dis. 2011; 58(5):729–736. [PubMed: 21872376]

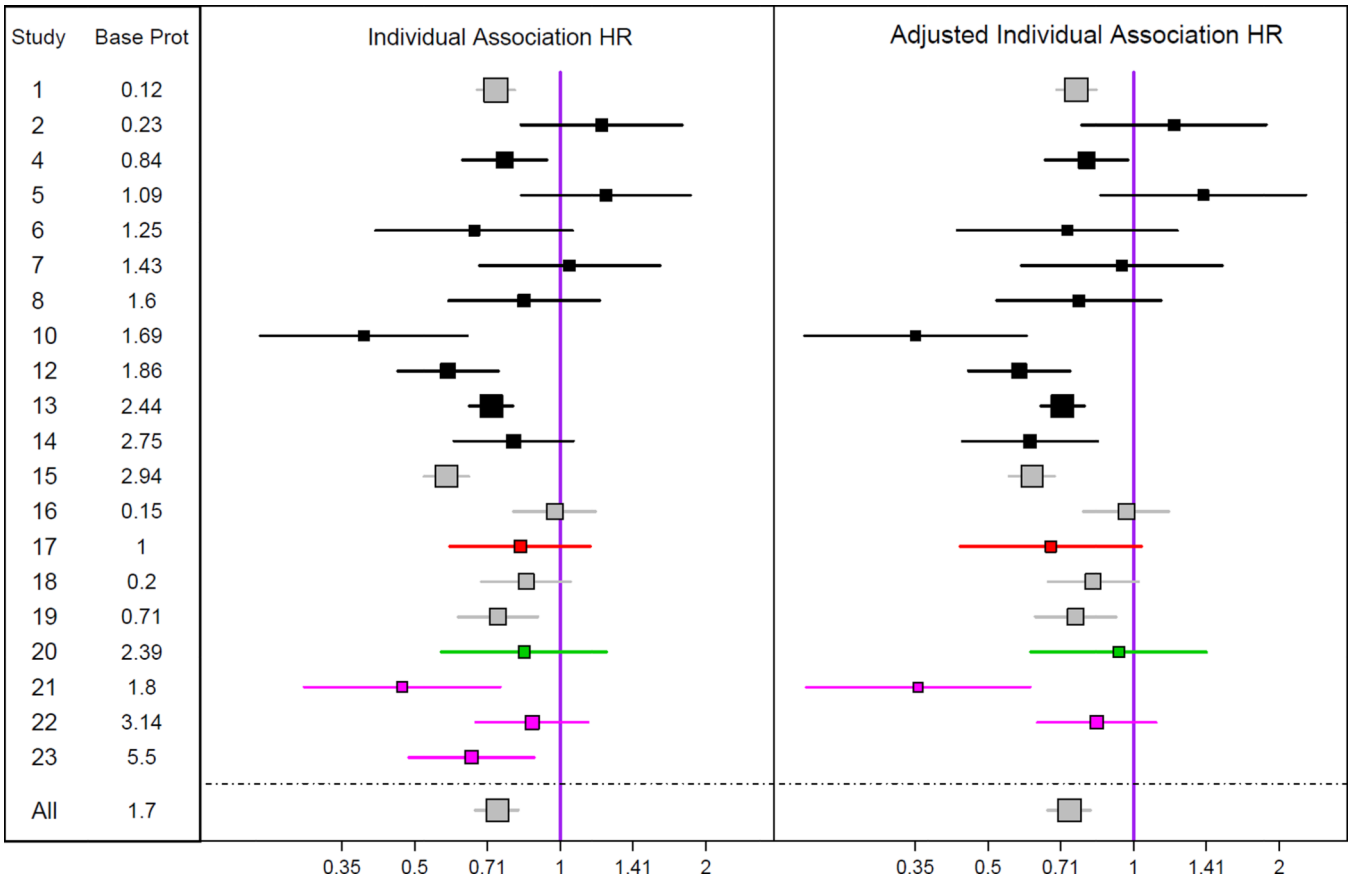

Figure 1. Individual Level Association

Shown are estimated hazard ratios and 95% confidence intervals relating the clinical outcome (time to doubling of serum creatinine, ESRD, or death) to early change in proteinuria, with adjustment for baseline proteinuria only (left) and with adjustment for baseline proteinuria, sex, age, baseline serum creatinine and baseline mean arterial pressure (right). Hazard ratios are expressed for 50% reduction in proteinuria. Analyses were performed for all studies, but results for individual studies are graphically displayed only for studies with greater than 20 events. Each study is represented only once. The bottom intervals are Bayesian credible intervals for the average results across all the trials. The colors indicate intervention type. Gray, studies which tested more than one intervention; black, renin-angiotensin system blockade vs. placebo; red, renin-angiotensin system blockade vs. calcium channel blocker, green, intensive blood pressure; magenta, immunosuppressive therapies.

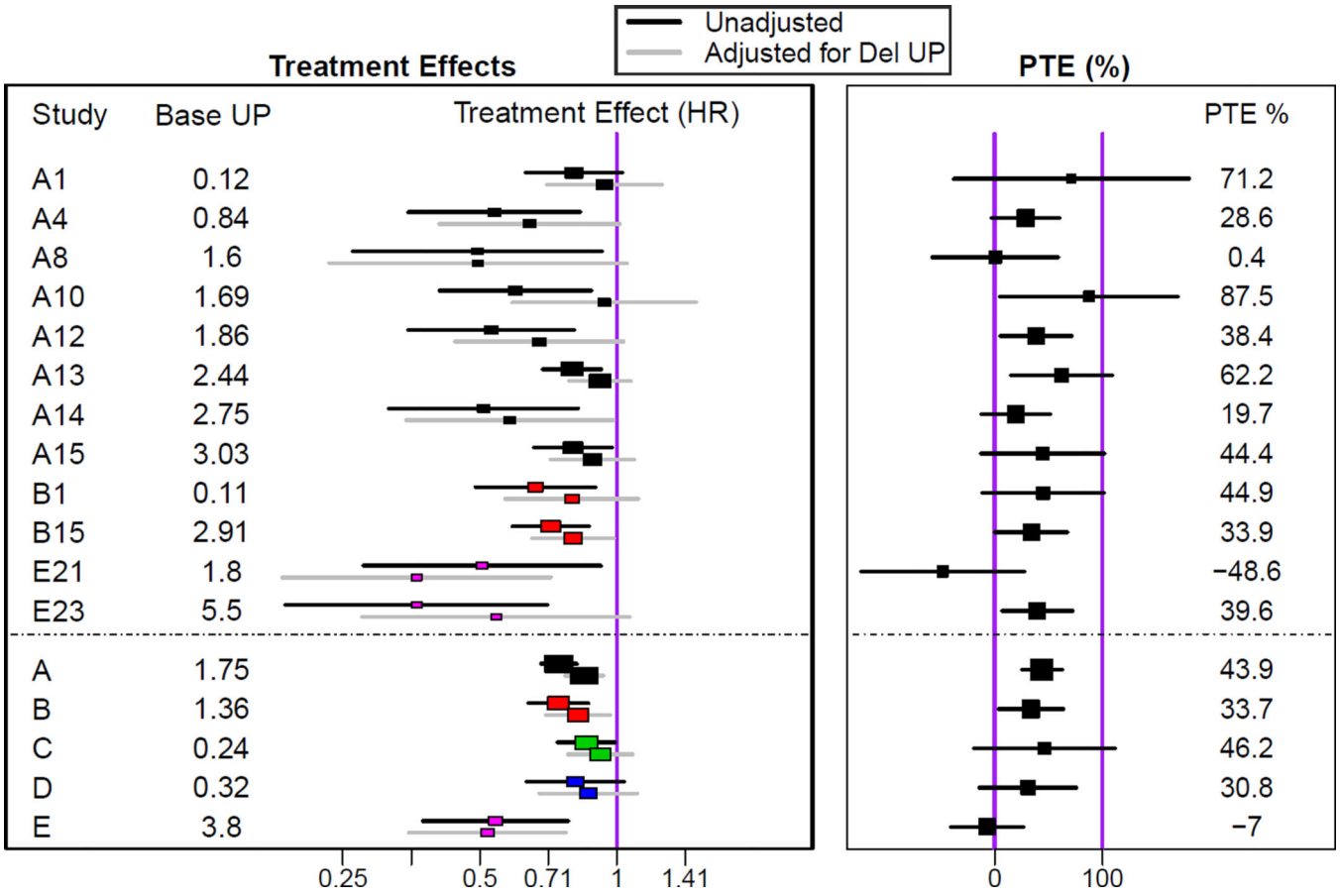

Figure 2. Evaluation of the Prentice-Freedman Criterion (Percent of Treatment Effect Explained)

Left: Shown are estimates of treatment effects on the clinical outcome (time to doubling of serum creatinine, ESRD, or death) without controlling for initial change in proteinuria (black) and then after controlling for early change in proteinuria (grey). Only studies with p < 0·10 in the unadjusted analysis are displayed. Both models controlled for baseline proteinuria, so the treatment effect estimates in this Figure differ from treatment effect estimates on the clinical outcome displayed in Figure 3. The bottom 5 intervals display pooled estimates of the unadjusted and adjusted treatment effects for the 5 treatment comparison classes under a fixed effect model. **Right:** Shown are estimates and associated 95% confidence intervals of the proportion of the treatment effect (PTE) explained by early change in proteinuria, defined as 1 minus the ratio of the adjusted to the unadjusted treatment effects (expressed on the log scale).

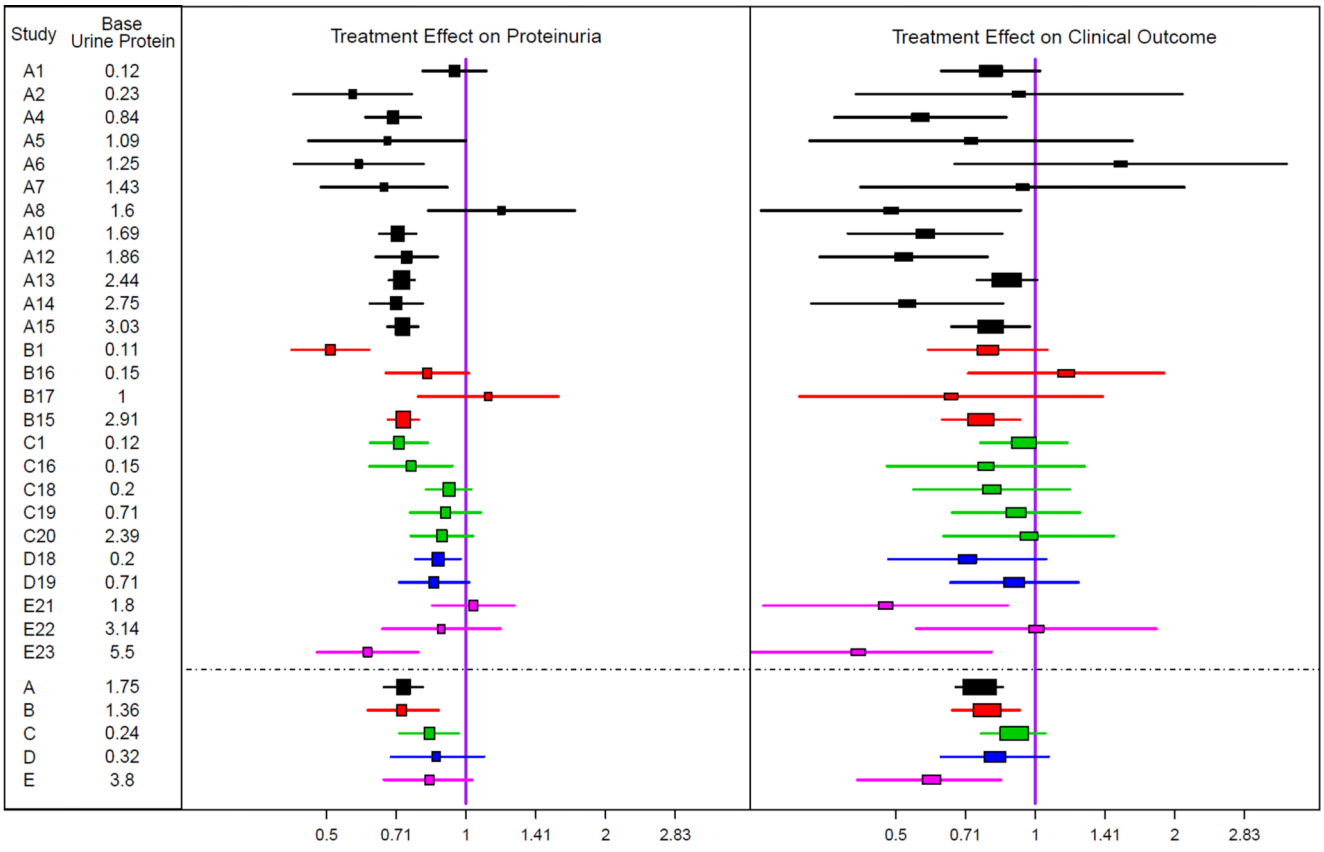

Figure 3. Treatment Effects on Change in Proteinuria and on the Clinical Outcome

Shown are geometric mean ratios and 95% confidence intervals comparing early change in proteinuria between treatment groups (left), and hazard ratios and 95% confidence intervals relating the clinical outcome (time to doubling of serum creatinine, ESRD, or death) to randomized treatment assignment (right). Analyses were performed for all studies, but results for individual studies are graphically displayed only for studies with greater than 20 events. Data for all studies is shown in Table S7. The bottom 5 intervals are Bayesian credible intervals for average results across the trials in the 5 treatment comparison classes under a random effects model. Proteinuria was log transformed.

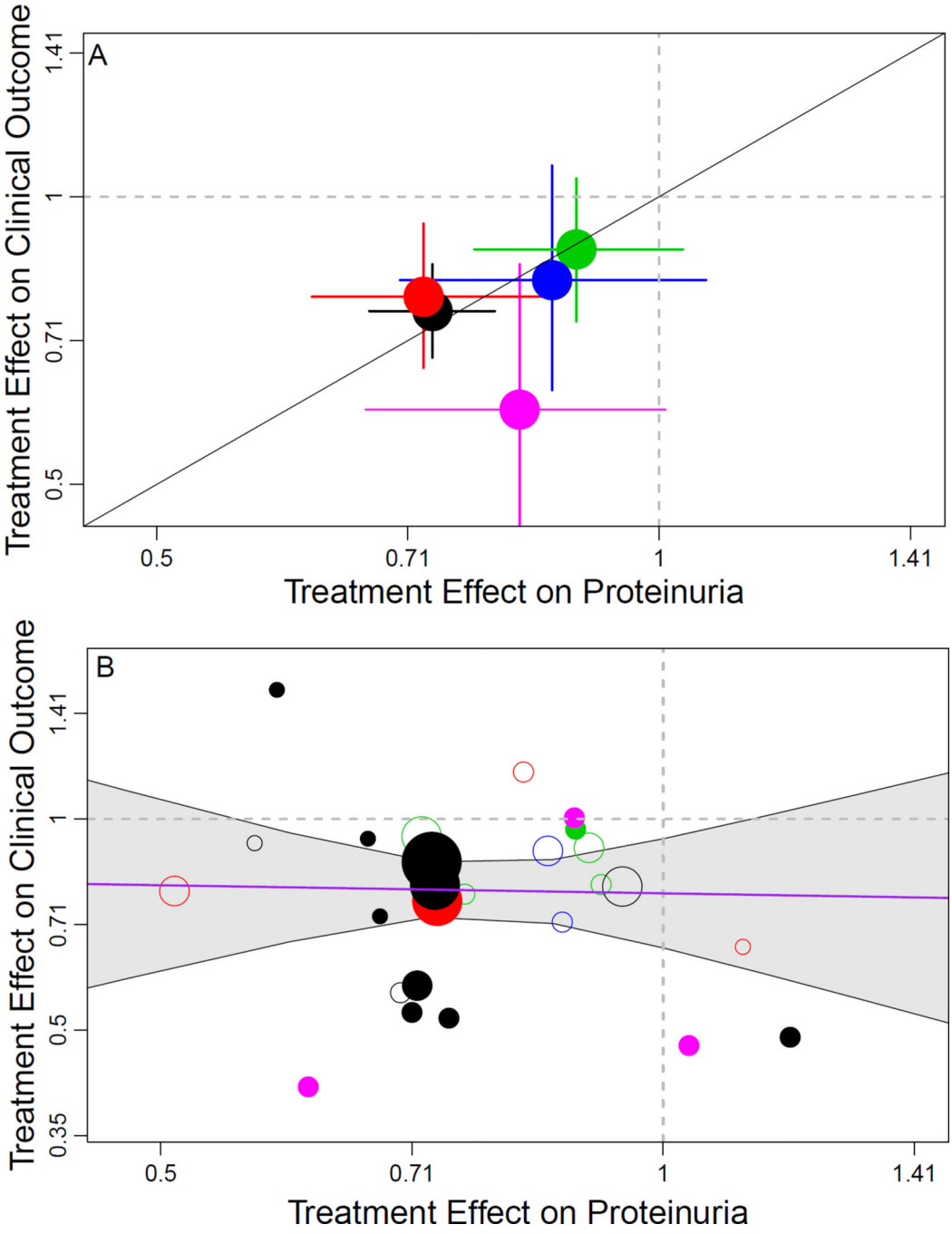

Figure 4. Trial-level Assessment of Validity of Proteinuria as a Surrogate Endpoint

Shown is the relationship between estimated treatment effects on the clinical outcome (time to doubling of serum creatinine, ESRD, or death) on the vertical axis to estimated treatment effects on the early change in proteinuria (on the horizontal axis). Treatment effects on the clinical outcome are expressed as hazard ratios and treatment effects on early change in proteinuria are expressed as geometric mean ratios. Proteinuria was log transformed in each analysis. The colors indicate intervention type. Black, renin-angiotensin system blockade vs. placebo; red, renin-angiotensin system blockade vs. calcium channel blocker, green,

intensive blood pressure; magenta, immunosuppressive therapies. **Top panel:** Aggregate results for the 5 interventions. The diagonal line is the line of identity. Solid lines extending from circle indicate the Bayesian credible intervals for the treatment effect on the clinical endpoint and change in urine protein. **Bottom panel:** Results for individual studies. The diameters of the circles are approximately proportional to the square root of numbers of events for the clinical outcome in each trial. The bolded circles indicate studies with median baseline proteinuria > 1 gram/day. The 95% Bayesian credible interval for the regression coefficient relating the treatment effects was too wide to be informative, ranging from a 0·60% lower to a 0·64% higher HR for the clinical endpoint for every 1% lower geometric mean ratio for proteinuria.

Table 1

Characteristics of Study and Study Groups

Study No.	Disease	N	Urine Protein (g/d)	Scr (mg/dL)	Events					Median F/U (y)
					ESRD	Doubling Scr	Death	Composite	%	
RAS Blockade vs Control										
A1^19	HTN	877	0·12 (004, 0·61)	1·99 (0·69)	135	134	83	258	29·4	4·2
A2^20	Mixed	122	0·23 (0·10, 0·98)	2·63 (1·42)	10	14	2	24	19·7	2·2
A3^21	Mixed	103	0·50 (0·10, 2·50)	1·82 (0·70)	7	10	4	14	13·6	3·1
A4^22	Mixed	562	0·84 (0·13, 2·46)	2·09 (0·62)	2	77	9	88	15·7	2·4
A5^23	Mixed	55	109 (0·44, 2·60)	5·05 (2·04)	21	9	2	25	46	2·2
A6^20	Mixed	106	1·25 (0·26, 3·14)	2·66 (1·21)	15	13	3	23	21·7	20
A7^24	Mixed	67	1·43 (0·75, 2·95)	4·25 (1·24)	15	11	2	24	36	1·3
A8^25	Mixed	98	1·60 (0·60, 3·20)	3·00 (0·74)	26	25	3	39	40	2·1
A9^26	IgA	109	1·60 (105, 2·61)	1·22 (0·50)	3	7	0	8	7·3	2·7
A10^27	Mixed	224	1·69 (106, 2·24)	3·94 (0·61)	83	47	1	109	48·7	2·2
A11^28	IgA	44	1·70 (1·10, 2·40)	100 (0·24)	15	6	0	15	34	7·8
A12^29	DM	409	1·86 (101, 3·84)	1·35 (0·44)	35	82	11	94	23·0	3·2
A13^30	DM	1513	2·44 (1·19, 5·18)	1·87 (0·48)	341	360	313	686	45·3	2·7
A14^31	Mixed	322	2·75 (1·53, 4·34)	2·20 (0·92)	58	40	4	73	22·7	2·1
A15^32	DM	1137	3·03 (1·76, 5·40)	1·69 (0·60)	182	231	175	405	35·6	2·6
RAS Blockade vs CCB s										
B1^19	HTN	653	0·11 (004, 0·57)	2·00 (0·71)	106	93	56	186	28·5	4·2
B16^33	DM	392	0·15 (007, 0·84)	1·17 (0·30)	0	24	46	66	16·8	4·6
B17^34	Mixed	121	100 (0·42, 2·60)	2·99 (1·01)	21	22	1	29	24·0	2·2
B15^32	DM	1129	2·91 (1·68, 5·11)	1·66 (0·58)	185	242	165	416	36·8	2·6
Intensive BP Control										
C1^19	HTN	1094	0·12 (0·04, 0·58)	2·00 (0·70)	179	164	105	328	30· 0	4·2
C16^33	DM	392	0·15 (0·07, 0·84)	1·17 (0·30)	0	24	46	66	16·8	4·6
C18^35	Mixed	584	0·20 (006, 1·12)	1·93 (0·52)	58	74	17	104	17·8	2·3

Study No.	Disease	N	Urine Protein (g/d)	Scr (mg/dL)	Events					Median F/U (y)
					ESRD	Doubling Scr	Death	Composite	%	
C19^35	Mixed	255	0·71 (0·17, 2·04)	3·43 (0·88)	136	63	17	154	60·4	2·3
C20^36	Mixed	330	2·39 (1·51, 3·66)	2·69 (1·11)	71	42	5	86	26·1	1·8
Low Prote in Diet										
D18^35	Mixed	584	0·20 (006, 1·12)	1·93 (0·52)	58	74	17	104	17·8	2·3
D19^35	Mixed	255	0·71 (0·17, 2·04)	3·43 (0·88)	136	63	17	154	60·4	2·3
Immunosuppressive therapy										
E21^37–40	IgA	233	1·80 (1·10, 3·31)	1·81 (0·75)	38	17	3	46	19·7	2·6
E22^41–43	Lupus	228	3·14 (1·91, 6·20)	1·51 (0·99)	31	22	21	46	20·2	4·8
E23^44–50	Membranous	343	5·50 (4·00, 8·30)	106 (0·34)	14	41	6	41	120	40
Pooled Analyses[a]										
A: RAS Blockade vs control		5748	1·75 (0·67, 3·71)	2·02 (0·93)	948	1066	612	1885	32·8	2·8
B: RAS Blockade vs. CCB		2295	1·36 (0·15, 3·23)	1·74 (0·73)	312	381	268	697	30·4	3·4
C: Intensive BP control		2655	0·24 (006, 1·44)	2·08 (0·92)	444	367	190	738	27·8	3·4
D: Low Protein Diet		839	0·32 (007, 1·51)	2·39 (0·95)	194	137	34	258	30·8	2·3
E: Immunosuppressive Therapy		804	3·80 (2·06, 6·30)	1·41 (0·77)	83	80	30	133	16·5	3·8
Baseline UP										
≤1 g/d		5393	0·14 (006, 0·50)	1·91 (0·85)	460	480	373	1034	19·2	3·6
>1–≤3 g/d		3648	1·81 (1·38, 2·35)	2·07 (102)	644	618	293	1151	31·6	2·9
>3 g/d		3281	501 (3·87, 7·34)	1·95 (0·87)	869	929	464	1514	46·1	2·5

Note: Unless otherwise indicated, values for categorical variables are given as number; values for continuous variables are given as mean ± standard deviation or median (25th, 75th quartiles). The Modification of Renal Diseases (MDRD) Study had recruited and analyzed by two strata of glomerular filtration rate (study A and study B), and are therefore listed as two studies. Abbreviations: BP, blood pressure; N, sample size; ESRD, end stage renal disease; Scr, serum creatinine; DM, diabetes mellitus; F/U, follow up; RAS; renin angiotensin system; HTN, hypertension; CKD, chronic kidney disease; IgA, immunoglobulin A; CCB, calcium channel blocker; BP, blood pressure;

[a]The total number of participants of the pooled analyses will be greater than the total number of participants since participants in studies with a factorial designs or with two intervention arms would be represented twice. Within each separate analysis, participants would have been included only once. The number of participants in baseline urine protein groups is less than the number of participants in the intervention groups due to 19 participants not having baseline urine protein.

Table 2

Dependence of Individual Level Association and Treatment Effects on Baseline Proteinuria

	Individual Level Association		Treatment Effect on ΔUP*(log scale)		Treatment Effect on ΔUP(raw scale, g/d)		Treatment Effect on Clinical Outcome	
UP Level (g/d)	% Reduction in Hazard[a]	95% Bayesian CI	% Reduction in GM[a]	95% Bayesian CI	ΔUP	95% Bayesian CI	% Reduction in Hazard[a]	95% Bayesian CI
<1	12	4 to 20	25	16 to 34	0.08	0.05–0.10	6	−8 to 18
1–3[b]	31	23 to 37	35	27 to 42	0.62	0.49–0.76	26	16 to 35
>3[b]	31	31 to 39	31	21 to 39	1.46	1.03–1.88	31	21 to 39

UP, urine protein; CI, credible interval; GM, geometric mean

[a]Hazard ratios and GM ratios have been converted to % reductions in hazards or % reductions in geometric means to facilitate comparisons with absolute treatment effects on change in proteinuria in g/d.

[b]Bayesian posterior probabilities exceed 0.975 indicating larger of the effects in the subgroups with baseline UP 1–3 g/day or > 3 g/day groups vs. the subgroup with baseline urine protein < 1 g/day group for each of the following analyses: 1) Individual level association, 2) Treatment effect on change in UP expressed on the raw scale (g/day) but not on the log scale (% reduction in GM), and 3) Treatment effect on the clinical outcome.