

Validation of Creatinine-Based Estimates of GFR When Evaluating Risk Factors in Longitudinal Studies of Kidney Disease

Xuelei Wang,* Julia Lewis,[†] Lawrence Appel,[‡] DeAnna Cheek,[§] Gabriel Contreras,^{||} Marquetta Faulkner,[¶] Harold Feldman,** Jennifer Gassman,* Janice Lea,^{††} Joel Kopple,^{‡‡} Mohammed Sika,[†] Robert Toto,^{§§} and Tom Greene,* for the AASK Investigators

*Cleveland Clinic Foundation, Quantitative Health Sciences, Cleveland, Ohio; [†]Nephrology Clinical Trials Center, Vanderbilt University Medical Center, Nashville, Tennessee; [‡]ProHealth, Johns Hopkins Medical Institutions, Baltimore, Maryland; [§]Pediatric Nephrology, Medical University of South Carolina, Charleston, South Carolina; ^{||}Division of Nephrology, University of Miami, Miami, Florida; [¶]Medicine/Research, Meharry Medical College, Nashville, Tennessee; **Center for Clinical Epidemiology & Biostatistics, University of Pennsylvania, Philadelphia, Pennsylvania; ^{††}Center for Hypertension & Renal Disease Research, Emory University, Atlanta, Georgia; ^{‡‡}Division of Nephrology and Hypertension, Harbor-UCLA Medical Center, Torrance, California; ^{§§}Division of Nephrology, University of Texas Southwestern Medical Center at Dallas, Dallas, Texas

Whereas much research has investigated equations for obtaining estimated GFR (eGFR) from serum creatinine in cross-sectional settings, little attention has been given to validating these equations as outcomes in longitudinal studies of chronic kidney disease. A common objective of chronic kidney disease studies is to identify risk factors for progression, characterized by slope (rate of change over time) or time to event (time until a designated decline in kidney function or ESRD). The relationships of 35 baseline factors with eGFR-based outcomes were compared with the relationships of the same factors with iothalamate GFR (iGFR)-based outcomes in the African American Study of Kidney Disease and Hypertension (AASK; $n = 1094$). With the use of the AASK equation to calculate eGFR, results were compared between time to halving of eGFR or ESRD and time to halving of iGFR or ESRD (with effect sizes expressed per 1 SD) and between eGFR and iGFR slopes starting 3 mo after randomization. The effects of the baseline factors were similar between the eGFR- and iGFR-based time-to-event outcomes (Pearson $R = 0.99$, concordance $R = 0.98$). Small but statistically significant differences ($P < 0.05$, without adjustment for multiple analyses) were observed for seven of the 35 factors. Agreement between eGFR and iGFR was somewhat weaker, although still relatively high for slope-based outcomes (Pearson $R = 0.93$, concordance $R = 0.92$). Effects of covariate adjustment for age, gender, baseline GFR, and urine proteinuria also were similar between the eGFR and iGFR outcomes. Sensitivity analyses including death in the composite time-to-event outcomes or using the Modification of Diet in Renal Disease equation instead of the AASK equation provided similar results. In conclusion, the data from the AASK provide tentative support for use of outcomes that are based on an established eGFR formula using serum creatinine as a surrogate for measured iGFR-based outcomes in analyses of risk factors for the progression of kidney disease.

J Am Soc Nephrol 17: 2900–2909, 2006. doi: 10.1681/ASN.2005101106

The GFR is regarded as the best overall index of renal function in patients with chronic kidney disease (CKD) (1). However, because measurement of GFR is expensive and logistically difficult, serum creatinine (SCr) often is used as an alternative (2,3). SCr is an imperfect indicator of GFR because it is influenced by creatinine generation and tubular secretion, both of which vary between individuals and within individuals over time. In an attempt to overcome this draw-

back, several equations have been developed to estimate GFR from SCr in conjunction with demographic factors (4–7). The most widely used is the Modification of Diet in Renal Disease (MDRD) equation, which was developed by applying linear regression to enrollees in the MDRD Study (5,6). A related equation was developed specifically for black individuals by applying similar methods to relate measured GFR, estimated by the clearance of I^{125} -iothalamate, to SCr, gender, and age at the baseline evaluation of the African American Study of Kidney Disease and Hypertension (AASK) (7).

To date, estimating equations for GFR have been derived and validated almost exclusively using cross-sectional data sets (4–18). Whereas limited work has been done to evaluate the association between longitudinal changes in creatinine-based estimates of GFR and contemporaneous longitudinal changes in

Received October 24, 2005. Accepted August 2, 2006.

Published online ahead of print. Publication date available at www.jasn.org.

Address correspondence to: Ms. Xuelei Wang, Cleveland Clinic Foundation, Department of Quantitative Health Sciences, Wb-4, 9500 Euclid Avenue, Cleveland, OH 44195. Phone: 216-445-4319; Fax: 216-445-2781; E-mail: wangx@ccf.org

iothalamate GFR (iGFR), including studies in lung transplant recipients (19) and Pima Indians with type 2 diabetes (20), the validity of creatinine-based outcomes to identify risk factors in longitudinal studies has not been examined comprehensively. Nonetheless, longitudinal changes in creatinine-based estimates of GFR are used routinely as outcomes in randomized, clinical trials and cohort studies (7,21–24).

We previously reported that despite subtle differences, the main conclusions of the randomized treatment group comparisons of the AASK were similar between outcomes that were based on estimated GFR (eGFR) using the AASK equation and outcomes that were based on iGFR (25). This report extends this work by examining the concordance of the relationships of 35 potential risk factors with eGFR-based outcomes *versus* the corresponding relationships of the same factors with iGFR-based outcomes. When iGFR is viewed as a reference standard, this analysis can be viewed as evaluating the validity of eGFR-based outcomes as surrogate end points in longitudinal studies in which the research objective is to identify risk factors for the progression of renal disease.

Both slope-based and time-to-event outcomes have been used in studies of renal disease progression. Slope-based analyses evaluate the average progression rate in all patients, whereas time-to-event analyses are more sensitive to large, clinically important declines in renal function (26). Because the suitability of eGFR as a surrogate for iGFR may differ between these approaches, we investigated the validity of eGFR for both slope and the time-to-event end points.

Materials and Methods

Patients and Renal Function Measurements

The AASK was a randomized clinical trial of black individuals who had hypertension ($n = 1094$), were aged 18 to 70 yr, and had a GFR between 20 and 65 ml/min per 1.73 m². Participants were randomly assigned according to a 3 × 2 factorial design to one of three antihypertensive drug regimens (first-line therapy with a calcium channel blocker [amlodipine], β blocker [metoprolol], or an angiotensin-converting enzyme inhibitor [ramipril]) and to two levels of BP control (mean arterial pressure ≤92 *versus* 102 to 107 mmHg). On the recommendation of the Data Safety and Monitoring Board, the amlodipine intervention was terminated approximately 1 yr before the end of the trial.

The GFR was assessed by renal clearance of ¹²⁵I-iothalamate twice at baseline, at 3 and 6 mo, then every 6 mo thereafter. For this report, iGFR was standardized to body surface area by multiplication by 1.73/(body surface area), where body surface area was computed using the patient's weight at the time of the GFR measurement. SCr was measured centrally using the rate-Jaffe method with an alkaline picrate assay (normal range 0.7 to 1.4 mg/dl) twice during baseline, at 3 and 6 mo, then at 6-mo intervals during follow-up. A total of 10,679 iGFR and 11,130 SCr measurements were obtained during the trial (excluding those in the amlodipine group after September 2000). To facilitate comparisons between analyses that were based on iGFR and SCr, the data set was restricted to 9742 matched pairs of iGFR and SCr measurements that were obtained within 8 wk of each other when evaluating time-to-event outcomes and to 8529 of these matched pairs that remained after exclusion of the second baseline measurement. Each of the SCr measurements was used to compute eGFR using the AASK equation (7) $eGFR = 329 \times SCr^{-1.096} \times age^{-0.294} \times (0.736 \text{ if female})$.

The mean follow-up period from randomization to the final iGFR–SCr pair was 3.6 yr. The protocol and procedures were approved by the institutional review board at each center, and all participants gave written informed consent. Additional details regarding the trial have been presented elsewhere (27–29).

Outcomes

For time-to-event analyses, we compared the effects of the baseline risk factors on the time from randomization to the two composite outcomes defined by (1) a 50% reduction in iGFR from the mean of two baseline values or ESRD or (2) a 50% reduction in eGFR from the mean of two baseline eGFR measurement or ESRD. For slope-based analyses, the statistical models evaluated the rate of change in iGFR and eGFR separately in the first 3 mo (acute phase) and the subsequent period after 3 mo (chronic phase) because the study interventions were known to lead to hemodynamic changes in renal function that differed from their hypothesized long-term effects. The data presentation of this report is limited to the chronic phase, which may reflect long-term disease progression better.

Potential Baseline Risk Factors

A total of 38 potential baseline variables were selected by the investigators before the analyses of the data for examination as potential risk factors for the progression of renal disease (30). Two of the selected factors, pulse pressure and mean arterial pressure, were excluded from this report because they are mathematical functions of systolic and diastolic BP, which also were selected. A third factor, baseline iGFR, was excluded because it is a component of the outcome variable in analyses of iGFR slope. The remaining 35 factors are listed in Table 1. For categorical factors, the reference group was defined by consensus of the investigators as the category that best represents the absence of the potential risk factor. With the exception of hematocrit, which was obtained locally, the remaining serum and urine biochemistry measurements were obtained at a central laboratory (Cleveland Clinic Research Laboratory, Cleveland, OH). BP measurements were taken as the average of two seated values by a random zero sphygmomanometer. Demographic information and medical histories were obtained by patient interview and chart review.

Statistical Analyses

Assessment of Agreement on a Patient Basis. The eGFR- and iGFR-based time-to-event outcomes were compared using two-way contingency tables and the κ statistic for time-to-event outcomes (31). The eGFR and iGFR slopes were compared for individual patients using the following measures:

1. Mean eGFR slope – mean iGFR slope (to assess bias)
2. Pearson correlation of eGFR slope with iGFR slope (to assess precision)
3. Concordance correlation of eGFR slope with iGFR slope (to assess agreement)
4. Root mean square error (rMSE) of iGFR slope that cannot be accounted for by a linear regression of iGFR slope on eGFR slope (to assess precision)

The concordance correlation (32) is a measure of agreement that adjusts the Pearson correlation downward if there is a systematic bias between the measures being compared. The rMSE is an estimate of the variability in the iGFR slopes that cannot be explained by the eGFR slopes. Measures 2 through 4 were adjusted for randomized treatment group. For patient-level analyses the chronic eGFR and iGFR slopes were computed using least squares regressions of each patient's eGFR

Table 1. Patient characteristics of 35 selected baseline risk factors^a

Baseline Factor (<i>n</i> = 1094 unless indicated otherwise)	Label for Figures	Mean ± SD or <i>n</i> (%)
Urine protein/creatinine (<i>n</i> = 1090)	up/cr	0.33 ± 0.52
Age at randomization (yr)	age	54.6 ± 10.7
Female gender	fem	424 (38.8)
Income <\$15,000	inc	521 (58.5)
Not a high school graduate (<i>n</i> = 1092)	edu	444 (40.7)
Current smoker	smoke	321 (41.0)
History of heart disease	hrt	564 (51.6)
Familial diagnosis of ESRD	fesrd	140 (12.8)
Parental diagnosis of hypertension	prhtn	756 (69.1)
Duration of hypertension (yr; <i>n</i> = 1084)	hyp	14.2 ± 10.1
Height (cm)	ht	171 ± 10
Body weight (kg)	wt	89.5 ± 20.7
Body mass index (kg/m ²)	bmi	30.6 ± 6.6
Systolic BP (mmHg)	sys	150 ± 24
Diastolic BP (mmHg)	dia	96 ± 14
Evidence of left ventricular hypertrophy	lvh	420 (38.4)
Abnormal electrocardiogram	ekg	880 (80.4)
Serum urea nitrogen (mg/dl)	sun	24.2 ± 9.7
Serum creatinine (mg/dl)	scr	2.00 ± 0.70
Hematocrit (%; <i>n</i> = 1092)	hct	39.4 ± 4.9
Albumin (g/dl)	alb	4.23 ± 0.32
Glucose (mg/dl)	glu	94.0 ± 15.8
Total cholesterol (mg/dl; <i>n</i> = 1084)	chol	212 ± 45
HDL (mg/dl; <i>n</i> = 1084)	hdl	48 ± 16
Total – HDL cholesterol (mg/dl; <i>n</i> = 1081)	ldl	163 ± 45
Triglycerides (mg/dl; <i>n</i> = 843)	trig	140 ± 81
Uric acid (mg/dl)	uric	8.3 ± 1.9
Calcium × phosphorus (mg ² /dl ²)	ca_p	31.9 ± 5.2
Phosphorus (mg/dl)	phos	3.52 ± 0.56
Urine total volume (L; <i>n</i> = 1093)	uvol	2.21 ± 0.94
Urine creatinine (g/d; <i>n</i> = 1090)	ucr	1.61 ± 0.64
Urine creatinine/wt (mg/d per kg; <i>n</i> = 1090)	ucrwt	18 ± 6
Urine urea nitrogen (g/d; <i>n</i> = 1093)	uun	8.29 ± 3.79
Urine sodium (g/d; <i>n</i> = 1093)	una	3.68 ± 1.98
Urine sodium/potassium (<i>n</i> = 1093)	unak	2.27 ± 1.19

^aData are means ± SD for continuous variables and *n* (%) for categorical variables. The percentage with income <\$15,000 was based on 891 participants who provided information. The percentage of current smokers was based on 782 participants who either were current smokers or never smoked.

or iGFR measurements *versus* follow-up time under a two-slope segmented model, with separate slopes in months 0 through 3 and in the remainder of follow-up.

Assessment of Agreement between Effects of Covariates. The 35 baseline risk factors were scaled to have unit SD to facilitate comparisons of effects for different regression coefficients. For the time-to-event outcomes, separate Cox proportional hazards analyses were applied to relate the eGFR and iGFR composite outcomes to each baseline factor, and robust sandwich variance-covariance estimators were used to account for the dependence of the outcomes in the same patient (33). For the slope-based outcomes, iGFR and eGFR were analyzed separately using the two-slope mixed-effects models (34,35) with fixed-effects terms to estimate the mean effects of baseline factors and randomized treatment group on the intercept and acute and chronic slopes and with random intercepts and acute and chronic slopes to represent deviations of individual patients around the group means. A two-band Toeplitz error structure was used to account for autocorrelation in neighboring

GFR measurements (36). Robust sandwich estimates were used to estimate SE and the correlation between the regression coefficients linking the eGFR and iGFR slopes to the respective baseline covariates (37).

These analytic methods first were used to relate the time-to-event and slope outcomes individually to each of the 35 standardized baseline covariates, with adjustment for randomized treatment assignment. The effects of the baseline factors on the eGFR and iGFR outcomes were compared by (1) plotting the regression coefficients (transformed to hazard ratios (HR) for the time-to-event outcomes) for the eGFR outcome *versus* the iGFR outcome, with the line of identity indicating perfect agreement; (2) presenting the Pearson and concordance correlations between the effects of the 35 covariates on the eGFR and iGFR outcomes; and (3) presenting the rMSE of the effects of the 35 covariates on iGFR that cannot be accounted for by a linear regression on the effects on eGFR. Because observational analyses usually are performed with adjustment for major nonmodifiable demographic factors and

previously identified risk factors, these analyses were repeated with gender, age, baseline iGFR, and baseline proteinuria (defined as the log-transformed urine protein-to-creatinine ratio) as covariates.

Several sensitivity analyses were conducted to evaluate the robustness of the results. First, because death is a competing risk for the occurrence of the renal events, the comparisons of the time-to-event outcomes were repeated after addition of death as a component of the eGFR and iGFR composites. Second, because SCr is a component of eGFR, summary statistics that evaluated the agreement of the eGFR- and iGFR-based outcomes were recomputed after deletion of baseline SCr from the baseline risk factors. Third, the analyses that compared the effects of the baseline factors on the eGFR and iGFR outcomes were repeated using the MDRD formula in place of the AASK equation to calculate eGFR. Fourth, the comparisons of the time-to-event outcomes were repeated for doubling of SCr in places of 50% reduction in GFR.

Results

Patient Characteristics at Baseline

The 35 potential baseline risk factors are summarized in Table 1. The mean (\pm SD) age was 54.6 ± 10.7 yr, and 61% were male. Average prestudy duration of hypertension was 14.2 ± 10.1 yr, and baseline systolic and diastolic BP were 150 ± 24 and 96 ± 14 mmHg, respectively. Mean GFR was 46 ± 13 ml/min per 1.73 m^2 .

Patient-Level Association for Time-to-Event and Slope Outcomes

Of 1094 AASK participants, 280 experienced a halving of iGFR or ESRD, and 240 experienced a halving of eGFR or ESRD. The two event outcomes agreed for 1020 (93%) of the participants ($\kappa = 0.81$). A total of 74 (7%) participants reached one outcome but not the other (Table 2). The greater rate of iGFR events compared with eGFR events may reflect, in part, a higher variability of iGFR measurements (3).

A total of 1012 patients with at least two follow-up GFR measurements were included in patient-level slope analyses. There were only minimal biases in the eGFR slope estimates (0.10 ml/min per $1.73 \text{ m}^2/\text{yr}$ for the chronic slopes). The Pearson R and concordance R between chronic eGFR and iGFR slopes were 0.62 and 0.60, respectively. Because of a greater precision of slopes that were estimated over a longer follow-up period, participants with >18 mo of follow-up time had a better patient-level agreement between chronic eGFR and iGFR slopes than did patients with shorter follow-up (Figure 1, Table 3).

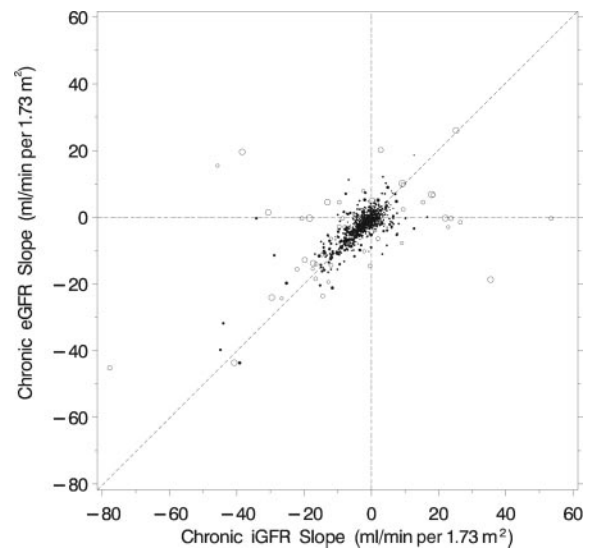


Figure 1. Agreement of chronic slopes on a patient basis. Shown is the association between the estimated GFR (eGFR) and the iothalamate GFR (iGFR) slopes during the chronic phase of the study starting 3 mo after randomization. Slopes are expressed in ml/min per $1.73 \text{ m}^2/\text{yr}$ and computed by the least squares method. The areas of the plot symbols are proportional to the SE of the slope estimates.

Association of Estimated Effects of Baseline Risk Factors without Covariate Adjustment

Figure 2 plots the effects of the 35 baseline factors on the eGFR time-to-event outcome versus the effects of the same factors on the iGFR-based time-to-event outcome. The abbreviations of each factor are defined in Table 1. The proximity of the plotted points to the line of identity indicates that the observed effects of the factors were similar between the iGFR- and eGFR-based outcomes, with Pearson $R = 0.99$ and concordance $R = 0.98$.

Twelve factors were statistically significant predictors of both the iGFR and eGFR time-to-event outcomes. Four factors had a significant effect on eGFR but not on iGFR; none had a significant effect on iGFR but not eGFR. Statistically significant (but generally small) differences between the two time-to-event outcomes were observed for seven of the 35 factors: Years of hypertension, age, diastolic BP, SCr, triglycerides, no high school education, and proteinuria. The largest discrepancies

Table 2. Patient-level association for time-to-event outcomes^a

iGFR-Based Outcome	eGFR-Based Outcome		
	Halving of eGFR or ESRD Event	No Event	Total
Halving of iGFR or ESRD event	223 (20.4%)	57 (5.2%)	280 (25.6%)
No event	17 (1.6%)	797 (72.9%)	814 (74.4%)
Total	240 (21.9%)	854 (78.1%)	1094 (100%)

^aShown is the relationship between the numbers (overall percentages) of patients who reached the composite end points of halving of estimated GFR (eGFR) or ESRD or of having of iothalamate GFR (iGFR) or ESRD during the follow-up period of the African American Study of Kidney Disease and Hypertension (AASK). The κ statistic relating the composite outcomes was 0.81.

Table 3. Patient-level association of chronic slopes between eGFR and iGFR^a

Parameter	Chronic Slope		
	Overall (n = 1012)	>18 Mo (n = 959)	≤18 Mo (n = 53)
Mean chronic iGFR slope (SD; ml/min per 1.73 m ² /yr)	−2.57 (6.91)	−2.41 (4.87)	−5.54 (21.91)
Mean chronic eGFR slope (SD; ml/min per 1.73 m ² /yr)	−2.47 (5.32)	−2.34 (4.40)	−4.86 (13.70)
Mean bias ^b (ml/min per 1.73 m ² /yr)	0.10	0.07	0.68
Pearson R	0.62	0.77	0.46
Concordance R	0.60	0.76	0.41
rMSE (ml/min per 1.73 m ² /yr) ^c	5.40	3.12	19.09

^aPatients with at least two follow-up GFR were included. Chronic slopes were computed by the least squares method. rMSE, root mean square error.

^bMean chronic eGFR slope – mean iGFR slope.

^crMSE of iGFR slope that cannot be accounted for by a linear regression of iGFR slope on eGFR slope.

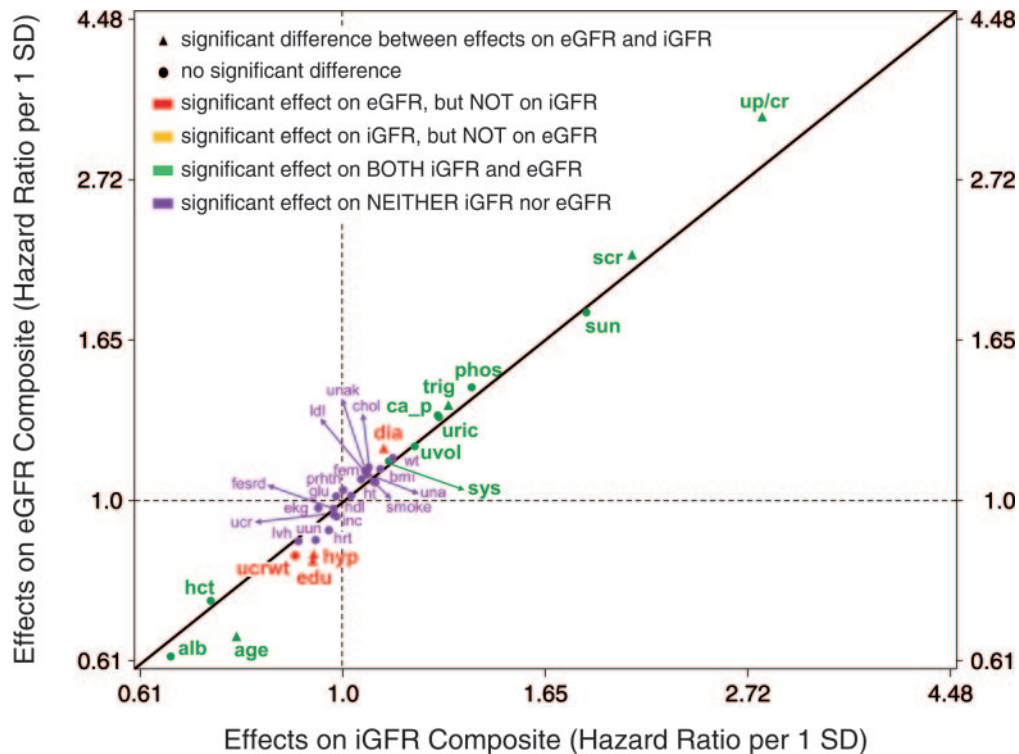


Figure 2. Agreement of effects of baseline factors on eGFR- and iGFR-based composite end points. Plotted are the hazard ratios (HR) of the eGFR composite outcome (vertical axis) versus the HR of the iGFR composite outcome (horizontal axis) associated with each of the 35 baseline factors. The line of identity indicates perfect agreement. HR are expressed per 1-SD increases in each factor. The color of the symbols indicates which factors were significantly related to the respective composite end points at the 5% level; triangles indicate factors whose effects on the eGFR and iGFR composites differ at the 5% level. Pearson R = 0.99, concordance R = 0.98, and root mean square error (rMSE) = 0.039 (log HR per 1 SD).

(with a >15% difference in the HR [per 1-SD increase]) occurred for proteinuria and age. Expressed in units that are relevant to each factor, doubling of proteinuria had an HR of 1.60 (95% confidence interval [CI] 1.51 to 1.70) for iGFR versus 1.73 (95% CI 1.62 to 1.84) for eGFR and a 10-yr increase in age with an HR of 0.78 (95% CI 0.70 to 0.87) for iGFR versus 0.67 (95% CI 0.60 to 0.75) for eGFR.

During the chronic phase, the Pearson and concordance R

relating the eGFR and iGFR slope outcomes were 0.93 and 0.92, respectively (Figure 3). Eight factors were observed to be significant predictors of both chronic slope outcomes, and four factors each had significant effects on eGFR but not on iGFR or iGFR but not eGFR. Six of the 35 factors had significantly different effects on the chronic iGFR and eGFR slopes, including SCr, diastolic BP, gender, height, urine volume, and triglyceride level. The largest discrepancy was observed for gender,

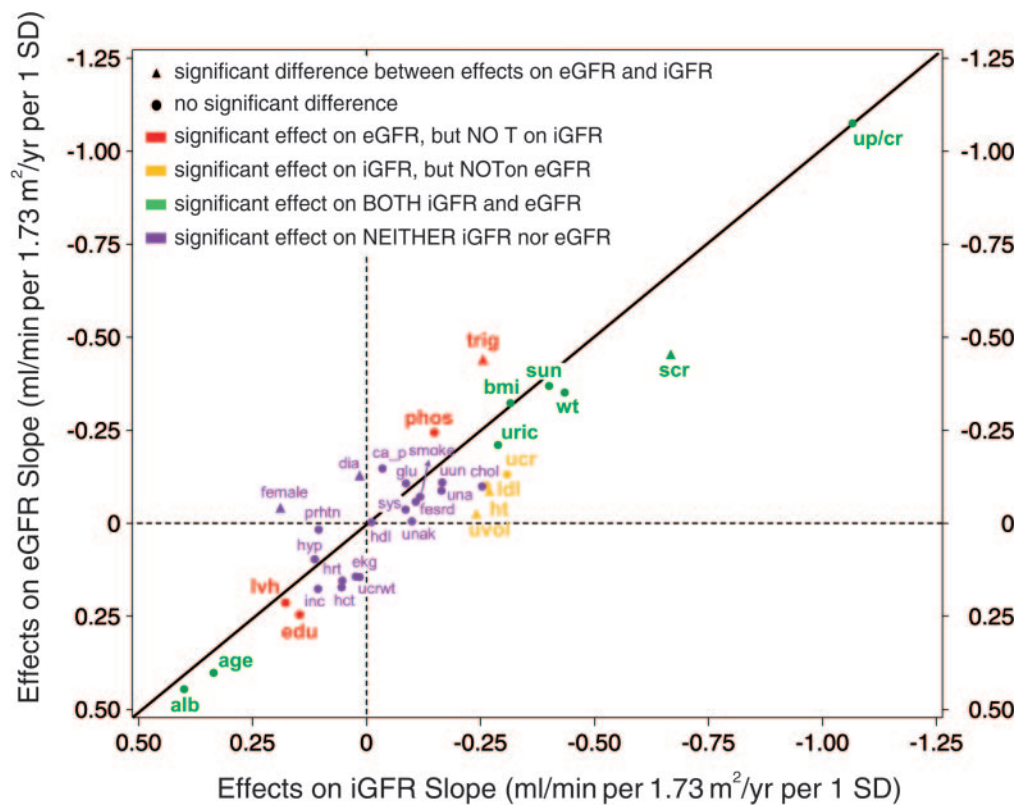


Figure 3. Agreement of effects of baseline factors on eGFR and iGFR slopes. Plotted are the effects of each of the 35 baseline risk factors on the chronic eGFR slope (vertical axis) and on the chronic iGFR slope (horizontal axis). Effects are expressed per 1-SD increases in the baseline factors. Plot symbols are defined as in Figure 2. Pearson $R = 0.93$, concordance $R = 0.92$, and $rMSE = 0.105$ ml/min per $1.73 \text{ m}^2/\text{yr}$ per 1 SD.

with female gender associated with a 0.40 ± 0.22 ml/min per $1.73 \text{ m}^2/\text{yr}$ steeper iGFR slope but a 0.087 ± 0.20 ml/min per $1.73 \text{ m}^2/\text{yr}$ less steep eGFR slope.

Effect of Covariate Adjustment on Agreement between the Effects of Baseline Factors on iGFR- and eGFR-Based Outcomes

Figure 4 plots the effects of the baseline factors on the eGFR-based time-to-event outcome *versus* the effects of the same factors on the iGFR-based time-to-event outcome without and with adjustment for age, gender, baseline GFR, and proteinuria. The interconnecting black lines describe the effect of covariate adjustment; lines that are parallel to the line of identity indicate similar effects of adjustment on the two outcomes. Covariate adjustment had similar effects on the two time-to-event outcomes, in both cases reducing the variation in estimated effects among the risk factors. As a result of this reduced variation, The Pearson and concordance R declined from 0.99 to 0.95 and from 0.98 to 0.93, respectively, after covariate adjustment. However, the stable $rMSE$ indicate that covariate adjustment did not reduce the precision with which estimates of effects on the iGFR composite could be estimated from the eGFR composite.

The effect of covariate adjustment was smaller for slope-based outcomes (Figure 5). With covariate adjustment, the Pearson R remained unchanged at 0.89, concordance R changed

from 0.87 to 0.86, and the $rMSE$ changed from 0.095 to 0.093 ml/min per $1.73 \text{ m}^2/\text{yr}$ per 1-SD change in the predictor variables.

Sensitivity Analyses

Of 1094 AASK participants, 354 had an event of halving of iGFR, ESRD, or death, and 317 had an event of halving of eGFR, ESRD, or death. The two event outcomes agreed for 1025 (94%) of the patients. The Pearson and concordance R of observed effects of the 35 factors relating the expanded iGFR and eGFR composites were both equal to 0.99.

Deleting SCr from the analyses had only minimal effects on the results, with relative changes in the Pearson and concordance R of <1% in all analyses.

The agreement of the eGFR-based outcomes with iGFR-based outcomes was similar when the MDRD equation was used in place of the AASK equation. When eGFR was computed by the MDRD equation, the Pearson and the concordance R between the effects of the 35 baseline factors on the eGFR and iGFR time-to-event composites were 0.99 and 0.98 and were 0.91 and 0.90, respectively, between the chronic eGFR and iGFR slopes. The Pearson and concordance R were 0.99 and 0.98, respectively, between the composite of doubling of SCr or ESRD and the composite of halving of iGFR or ESRD.

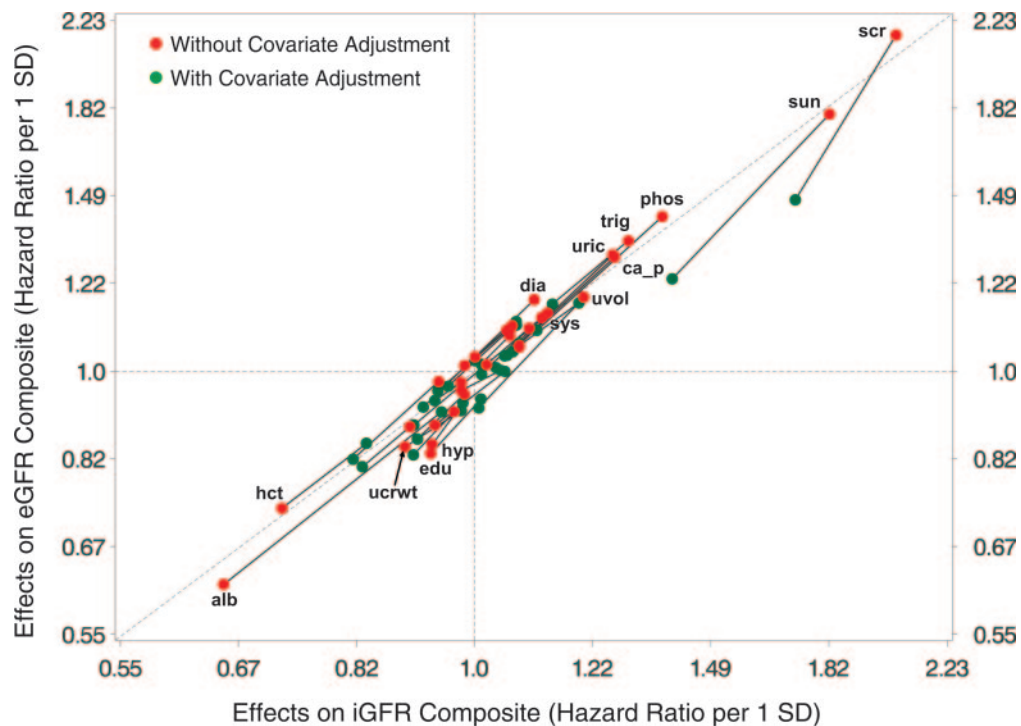


Figure 4. Effect of covariate adjustment on estimated effects of baseline factors on eGFR- and iGFR-based composite end points. Plotted are the HR of the eGFR composite outcome (vertical axis) versus the HR of the iGFR composite outcome (horizontal axis) associated with each of 32 baseline factors. HR are expressed per 1-SD increases in the indicated baseline factors. HR are displayed without (red circles) and with (green circles) adjustment for age, gender, baseline GFR, and baseline proteinuria. The solid black lines indicate the effect of covariate adjustment on the HR associated with the two outcomes. The Pearson and concordance R declined from 0.99 to 0.95 and from 0.98 to 0.93, respectively, after covariate adjustment. rMSE changed from 0.036 to 0.045 (log HR per 1 SD). Labels are provided for factors significantly associated ($P < 0.05$) with one or both outcomes without covariate adjustment.

Discussion

Using the AASK equation (7) to obtain eGFR from SCr, the estimated effects of 35 potential baseline risk factors on eGFR-based outcomes were found to have good overall agreement with the estimated effects of the same factors on iGFR-based outcomes. This was especially true for time-to-event outcomes, defined as composites of ESRD with a 50% reduction in either iGFR or eGFR (Figure 2). The Pearson and concordance correlations relating the effects of the 35 factors on the eGFR and iGFR composites were equal to 0.99 and 0.98, indicating that 96 to 98% of the variance of the effects of the 35 factors on iGFR could be accounted for by the effects of these factors on eGFR. Both outcomes identified the same factors as the strongest predictors of faster progression (higher urine proteinuria, SCr, serum urea nitrogen, serum phosphorus, and serum triglycerides) and the same predictors of slower progression (higher serum albumin and serum hematocrit and greater age). Agreement was somewhat weaker, although still relatively high, for the mean slope after 3 mo, with Pearson and concordance correlations of 0.93 and 0.92, respectively. For both slope-based outcomes, higher proteinuria and SCr were the strongest predictors of faster progression, and higher serum albumin and greater age were the strongest indicators of slower progression.

The identification of proteinuria as the strongest predictor of

progression for both eGFR- and iGFR-based outcomes is consistent with previous reports from the AASK describing the dominant role of proteinuria in predicting progression despite a low baseline median urine protein/creatinine ratio of 0.08 (30,38). The inverse relationship of age with progression may reflect that for a given entry GFR, older patients, on average, may have had their disease progress over a longer time period, in which case younger age would act as a marker for more rapidly progressive disease.

Although there generally was good agreement between the eGFR- and iGFR-based outcomes, discordant results were observed in specific cases. At the 5% significance level, seven of the 35 factors exhibited differences in their effects on the eGFR and iGFR time-to-event outcomes, and six factors exhibited differences for the slope outcomes. As shown in Figures 2 and 3, most of these differences were small compared with the magnitudes of the treatment effects, and some may be spurious as a result of the number of factors considered.

The predictor variables of commonly used equations for estimation of GFR include a term approximating the reciprocal of SCr ($\text{SCr}^{-1.096}$ for the AASK equation), age, gender, and, in some cases, race and weight. Among these factors, race and gender are constant over time, and in adults, changes in age and weight usually are too small during a follow-up period of several years to have a large effect on the computed eGFR.

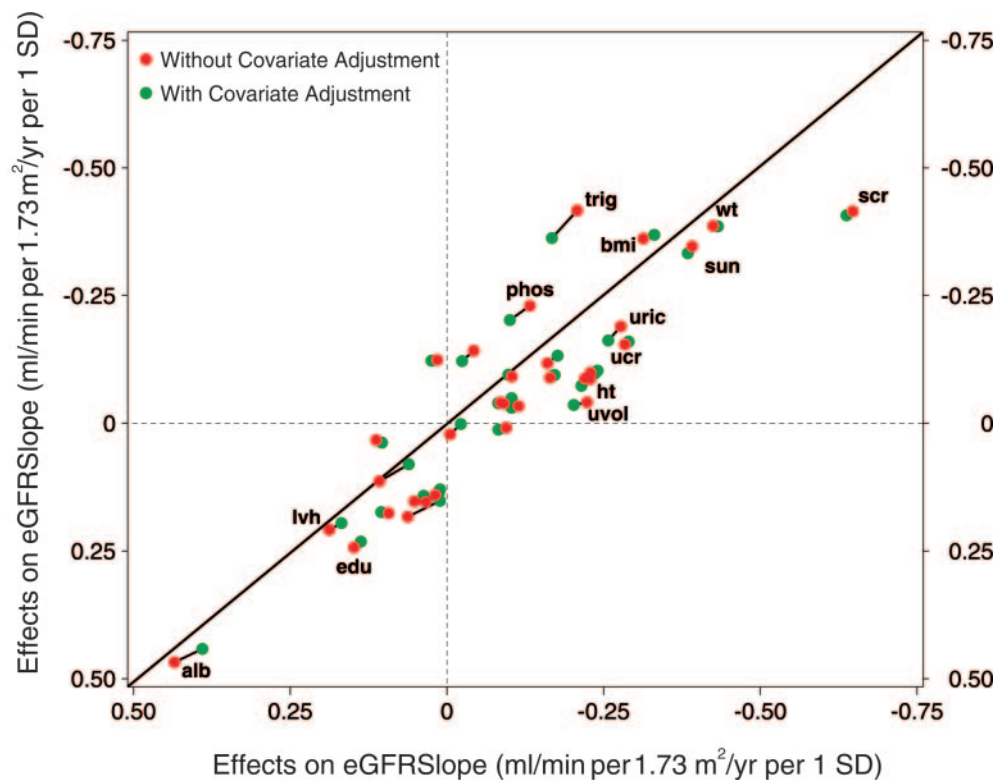


Figure 5. Effect of covariate adjustment on estimated effects of baseline factors on eGFR and iGFR slopes. Plotted are the effects of each of the 32 baseline risk factors on the chronic eGFR slope (vertical axis) and on the chronic iGFR slope (horizontal axis). Effects are expressed per 1-SD increases in the baseline factors. Estimated effects are displayed without (red circles) and with (green circles) adjustment for age, gender, baseline GFR, and baseline proteinuria. The solid black lines indicate the effect of covariate adjustment. With covariate adjustment, Pearson R changed from 0.89 to 0.89, concordance R changed from 0.87 to 0.86, and rMSE changed from 0.095 to 0.093 ml/min per 1.73 m²/yr per 1 SD. Labels are provided for factors significantly associated ($P < 0.05$) with one or both outcomes without covariate adjustment.

Hence, when assessing longitudinal change, similar results would be expected irrespective of which equation is used. In sensitivity analyses, the estimated effects of the 35 baseline factors on the eGFR slope and time-to-event outcomes were similar for the AASK and MDRD equations. In addition, the estimated effects on doubling of SCr, which often is used as an end point in randomized, clinical trials, were similar to the estimated effects on the halving of eGFR with the AASK equation. Hence, the results of this report apply also to time-to-event outcomes that are based on the doubling of SCr and to other equations for eGFR that stipulate an approximately reciprocal relationship of GFR with SCr.

Although the focus of this article is the agreement between results that are given by eGFR- and iGFR-based outcomes, it also is of interest to compare results between the slope-based and time-to-event outcomes. Some differences are notable; for example, a comparison of Figures 2 and 3 indicates that higher values of baseline weight and baseline body mass index were associated with steeper iGFR and eGFR slope but were not significantly associated with the iGFR or eGFR time-to-event outcomes.

Because the effects of potential risk factors usually are evaluated in observational studies after adjustment for key covariates, we also examined the effect of covariate adjustment for age, gender, and baseline proteinuria and GFR. Adjustment for

these factors reduced the variation in the estimated effects of the remaining factors on both the eGFR and iGFR time-to-event outcomes (Figure 4). However, the changes in the estimated effects that resulted from covariate adjustment were similar for the eGFR- and iGFR-based slope outcomes. In contrast to the time-to-event analyses, covariate adjustment resulted in only minimal changes in the estimated effects on GFR slope (Figure 5). Covariate adjustment seems to have had a greater effect on time-to-event than slope-based outcomes because a number of the baseline factors were strongly associated with the baseline GFR, which was more strongly related to the time-to-event outcomes than to the slope outcomes.

The analytic criteria that were used in this report are in some respects analogous to individual- and trial-level criteria that were proposed recently in the statistical literature on the validation of surrogate end points for randomized, clinical trials (39–41). An intermediate end point is regarded as a good surrogate on the *individual level* when the surrogate is strongly associated with the true clinical end point for individual patients and on the *trial level* when the treatment effects on the true end point can be predicted accurately from the treatment effects on the intermediate end point. Similarly, the analyses of Figure 1 and Table 2 examine the validity of eGFR-based outcomes as surrogates for iGFR-based outcomes at the individual

level. The analyses of Figures 2 through 5 evaluate whether effects of baseline factors on the iGFR-based outcomes can be predicted accurately from the effects of the same factors in the eGFR-based outcomes and are analogous to the trial-level criterion in that both approaches evaluate whether effects on the target clinical end point can be predicted accurately from effects on the surrogate.

The consideration of large numbers of potential risk factors in the longitudinal cohort setting increases the burden of proof for validation of a surrogate end point, because one has to consider, for each factor, whether that factor may affect the surrogate independent of the target outcome. In this study, we have addressed this issue by including a wide range of potential risk factors that cover as many domains as possible in our assessment of validity. Nonetheless, it is impossible to eliminate completely the risk that new potential risk factors that were not included in previous validation studies may behave differently with respect to the difference between eGFR and iGFR than those considered in previous validation studies.

There are several limitations in the scope of our analyses. First, because relationships of potential risk factors with progression may differ between populations, the results from this report of a single study should be interpreted cautiously until confirmed in other studies. Second, we have considered the validity of eGFR-based outcomes for identification of *baseline* risk factors that are measured before the follow-up period; it remains to be seen whether eGFR performs similarly well for evaluation of the effects of changes in risk during the follow-up period. Third, although the iothalamate-based estimate of GFR used as the target end point in this study is regarded as a rigorous method for estimating GFR in practice, iothalamate clearance may differ from inulin clearance, which is regarded as a truer “gold standard.” It has been estimated that up to 10% of iothalamate is secreted, and it is possible that new risk factors could alter iGFR by altering iothalamate handling by the kidney. We have not addressed in this article the association of iGFR- and eGFR-based outcomes with the occurrence of renal failure, which is the clinical end point of greatest interest. Fourth, we have limited the slope-based analyses to GFR itself, without log transformation. The log transformation expresses change in GFR on a percentage basis, which accords more closely with the time-to-event outcomes that are based on halving of GFR. However, because mean GFR declined faster at lower GFR levels in the AASK (25), the log transformation introduces other analytic complications that are beyond the scope of this article. Finally, the measures of agreement that are presented in association with Figures 2 through 5 pertain to the *estimated* effects of the baseline factors on eGFR- and iGFR-based outcomes that were observed in the study cohort rather than to the *true* effects in the target population. Methods for evaluating the relationships among the true effects are the subject of a separate statistical manuscript in preparation.

Conclusion

The data from the AASK provide tentative support for the use of outcomes that are based on eGFR, as determined by standard equations from SCr, as surrogates for corresponding

outcomes that are based on measured GFR for analyses of risk factors in longitudinal studies of the progression of kidney disease. As with other applications of surrogate end points, the use of eGFR-based outcomes to investigate new risk factors that have not been previously studied requires an extrapolation beyond existing data and should be done with appropriate caution.

References

1. Smith HW: Diseases of the kidney and urinary tract. In: *The Kidney: Structure and Function in Health and Disease*, New York, Oxford University Press, 1951, pp 836–887
2. Levey AS: Measurement of renal function in chronic renal disease. *Kidney Int* 38: 167–184, 1990
3. Perrone RD, Madias NE, Levey AS: Serum creatinine as an index of renal function: New insights into old concepts. *Clin Chem* 38: 1933–1953, 1992
4. Cockcroft DW, Gault MH: Prediction of creatinine clearance from serum creatinine. *Nephron* 16: 31–41, 1976
5. Levey AS, Bosch JP, Lewis JB, Greene T, Rogers N, Roth D; the Modification of Diet in Renal Disease Study Group: A more accurate method to estimate glomerular filtration rate from serum creatinine: A new prediction equation. *Ann Intern Med* 130: 461–470, 1999
6. Levey AS, Greene T, Kusek JW, Beck GJ: A simplified equation to predict glomerular filtration rate from serum creatinine [Abstract]. *J Am Soc Nephrol* 11: 155A, 2000
7. Lewis J, Agodoa L, Cheek D, Greene T, Middleton J, O'Connor D, Ojo A, Phillips R, Sika M, Wright J: Comparison of cross-sectional renal function measurements in African Americans with hypertensive nephrosclerosis and of primary formulas to estimate glomerular filtration rate. *Am J Kidney Dis* 38: 744–753, 2001
8. Vervoort G, Willems H, Wetzels J: Assessment of glomerular filtration rate in healthy subjects and normoalbuminuric diabetic patients: Validity of a new (MDRD) prediction equation. *Nephrol Dial Transplant* 17: 1909–1913, 2002
9. Bostom AG, Kronenberg F, Ritz E: Predictive performance of renal function equations for patients with chronic kidney disease and normal serum creatinine levels. *J Am Soc Nephrol* 13: 2140–2144, 2002
10. Skluzacek P, Szewc RG, Nolan CR 3rd, Riley DJ, Lee S, Pergola PE: Prediction of GFR in liver transplant candidates. *Am J Kidney Dis* 42: 1169–1176, 2003
11. Lin J, Knight EL, Hogan ML, Singh AK: A comparison of prediction equations for estimating glomerular filtration rate in adults without kidney disease. *J Am Soc Nephrol* 14: 2573–2580, 2003
12. Rule AD, Larson TS, Bergstralh EJ, Slezak JM, Jacobsen SJ, Cosio FG: Using serum creatinine to estimate glomerular filtration rate: Accuracy in good health and in chronic kidney disease. *Ann Intern Med* 141: 929–937, 2004
13. Hallan S, Asberg A, Lindberg M, Johnsen H: Validation of the Modification of Diet in Renal Disease formula for estimating GFR with special emphasis on calibration of the serum creatinine assay. *Am J Kidney Dis* 44: 84–93, 2004
14. Gaspari F, Ferrari S, Stucchi N, Centemeri E, Carrara F, Pellegrino M, Gherardi G, Gotti E, Segoloni G, Salvadori M, Rigotti P, Valente U, Donati D, Sandrini S, Sparacino V, Remuzzi G, Perico N; MY.S.S. Study Investigators: Performance of different prediction equations for estimating re-

- nal function in kidney transplantation. *Am J Transplant* 4: 1826–1835, 2004
15. Poggio ED, Wang X, Greene T, Van Lente F, Hall PM: Performance of the modification of diet in renal disease and Cockcroft-Gault equations in the estimation of GFR in health and in chronic kidney disease. *J Am Soc Nephrol* 16: 459–466, 2005
 16. Froissart M, Rossert J, Jacquot C, Paillard M, Houillier P: Predictive performance of the MDRD and Cockcroft-Gault equations for estimating renal function. *J Am Soc Nephrol* 16: 763–773, 2005
 17. Zuo L, Ma YC, Zhou YH, Wang M, Xu GB, Wang HY: Application of glomerular filtration rate estimating equations in Chinese patients with chronic kidney disease. *Am J Kidney Dis* 45: 463–472, 2005
 18. Rigalleau V, Lasseur C, Perlemoine C, Barthe N, Raffaitin C, Liu C, Chauveau P, Baillet-Blanco L, Beauvieux MC, Combe C, Gin H: Estimation of glomerular filtration rate in diabetic subjects: Cockcroft formula or Modification of Diet in Renal Disease study equation? *Diabetes Care* 28: 838–843, 2005
 19. Broekroelofs J, Stegeman CA, Navis GJ, de Haan J, van der Bij W, de Boer WJ, de Zeeuw D, de Jong PE: Creatine-based estimation of rate of long term renal function loss in lung transplant recipients. Which method is preferable? *J Heart Lung Transplant* 19: 256–262, 2000
 20. Perkins AB, Nelson GR, Ostrander EP, Blouch LK, Krolewski SA, Myers DB, Warram HJ: Detection of renal function decline in patients with diabetes and normal or elevated GFR by serial measurements of serum cystatin C concentration: Results of a 4-year follow-up study. *J Am Soc Nephrol* 16: 1404–1412, 2005
 21. Hansson L, Lindholm LH, Niskanen L, Lanke J, Hedner T, Niklason A, Luomanmaki K, Dahlof B, de Faire U, Morlin C, Karlberg BE, Wester PO, Bjorck JE: Effect of angiotensin-converting-enzyme inhibition compared with conventional therapy on cardiovascular morbidity and mortality in hypertension: The Captopril Prevention Project (CAPPP) randomised trial. *Lancet* 353: 611–616, 1999
 22. Brenner BM, Cooper ME, de Zeeuw D, Keane WF, Mitch WE, Parving HH, Remuzzi G, Snapinn SM, Zhang Z, Shahinfar S; RENAAL Study Investigators: Effects of losartan on renal and cardiovascular outcomes in patients with type 2 diabetes and nephropathy. *N Engl J Med* 345: 861–869, 2001
 23. Lewis EJ, Hunsicker LG, Clarke WR, Berl T, Pohl MA, Lewis JB, Ritz E, Atkins RC, Rohde R, Raz I; Collaborative Study Group: Renoprotective effect of the angiotensin receptor antagonist irbesartan in patients with nephropathy due to type 2 diabetes. *N Engl J Med* 345: 851–860, 2001
 24. Appel L, Agodoa L, Bakris G, Charleston J, Douglas J, Gassman J, Greene T, Jamerson K, Kusek J, Lewis JA, Middleton J, Miller ER, Lipkowitz M, Norris K, Phillips R, Rostand SG, Wright JT: The rationale and design of the AASK Cohort Study (AASK). *J Am Soc Nephrol* 14[Suppl 2]: S166–S172, 2003
 25. Lewis J, Greene T, Appel L, Contreras G, Douglas J, Lash J, Toto R, Van Lente F, Wang X, Wright JT Jr: A comparison of iothalamate-GFR and serum creatinine-based outcomes: Acceleration in the rate of GFR decline in the African American Study of Kidney Disease and Hypertension. *J Am Soc Nephrol* 15: 3175–3183, 2004
 26. Greene T, Lai J, Levey AS: Interpretation of clinical studies of renal disease. In: *Immunologic Renal Disease*, 2nd Ed., edited by Neilson EG, Couser WG, Philadelphia, Lippincott-Raven Publishers, 2001, pp 887–914
 27. Gassman J, Greene T, Wright JT Jr, Agodoa L, Bakris G, Beck GJ, Douglas J, Jamerson K, Lewis J, Kutner M, Randall OS, Wang SR: Design and statistical aspects of the African American Study of Kidney Disease of Hypertension (AASK). *J Am Soc Nephrol* 14[Suppl 2]: S154–S165, 2003
 28. Agodoa L, Bakris G, Wright J, Greene T, Beck G, Bourgoignie J, Briggs JP, Charleston J, Cheek D, Cleveland W, Douglas JG, Douglas M, Dowie D, Faulkner M, Gabriel A, Gassman J, Greene T, Hall Y, Hebert L, Hiremath L, Jamerson K, Johnson CJ, Kopple J, Kusek J, Lash J, Lea J, Lewis JB, Lipkowitz M, Massry S, Middleton J, Miller ER 3rd, Norris K, O'Connor D, Ojo A, Phillips RA, Pogue V, Rahman M, Randall OS, Rostand S, Schulman G, Smith W, Thornley-Brown D, Tisher CC, Toto RD, Wright JT Jr, Xu S; African American Study of Kidney Disease and Hypertension (AASK) Study Group: Effect of ACE inhibitor vs. dihydropyridine calcium antagonist-based treatment on renal outcomes in hypertensive nephrosclerosis. *JAMA* 285: 2719–2728, 2001
 29. Wright J, Agodoa L, Greene T, Agodoa LY, Appel LJ, Charleston J, Cheek D, Douglas-Baltimore JG, Gassman J, Glasscock R, Hebert L, Jamerson K, Lewis J, Phillips RA, Toto RD, Middleton JP, Rostand SG; African American Study of Kidney Disease and Hypertension Study Group: Effect of blood pressure lowering and antihypertensive drug class on progression of hypertensive kidney disease: Results from the AASK trial. *JAMA* 288: 2421–2431, 2002
 30. Baseline predictors of renal disease progression in the African American Study of Hypertension and Kidney Disease. *J Am Soc Nephrol* 17: 2928–2936, 2006
 31. Cohen J: A coefficient of agreement for nominal data. *Educ Psychol Meas* 20: 37–46, 1960
 32. Lin L: A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 45: 255–268, 1989
 33. Therneau T, Grambsch MP: *Modeling Survival Data*, New York, Springer, 2000, pp 169–229
 34. Verbeke G, Molenberghs G: *Linear Mixed Models for Longitudinal Data*, New York, Springer, 2001, p 24.3
 35. Laird NM, Ware JH: Random effects models for longitudinal data. *Biometrics* 38: 963–974, 1982
 36. *SAS/STAT User's Guide*, Version 8, SAS Publishing, North Carolina, 1999, pp 2137–2138
 37. Wei LJ, Lin DY, Weissfeld L: Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *J Am Stat Assoc* 84: 1065–1073, 1989
 38. Lea J, Greene T, Hebert L, Lipkowitz M, Massry S, Middleton J, Rostand S, Miller EP, Smith W, Bakris GL; for the AASK Study Investigators: Magnitude of proteinuria reduction predicts risk of end-stage renal disease: Results of the AASK trial. *Arch Intern Med* 165: 947–953, 2005
 39. Daniels MJ, Hughes MD: Meta-analysis for the evaluation of potential surrogate markers. *Stat Med* 16: 1515–1527, 1997
 40. Buyse M, Molenberghs G, Burzykowski T, Renard D, Geys H: The validation of surrogate endpoints in meta-analyses of randomized experiments. *Biostatistics* 1: 49–67, 2000
 41. Molenberghs G, Buyse M, Burzykowski T: A meta-analytic validation framework for continuous outcomes. In: *The Evaluation of Surrogate Endpoints*, edited by Burzykowski T, Molenberghs G, Buyse M, New York, Springer, 2005, pp 96–120