Accuracy of a Glomerular Filtration Rate Estimating Equation over Time in People with a Wide Range of Kidney Function

A thesis

submitted by

Smita Padala, M. D.

In partial fulfillment of the requirements
for the degree of

Master of Science

in

Clinical Research

TUFTS UNIVERSITY

Sackler School of Graduate Biomedical Sciences

May 2010

Mentors:
Andrew S. Levey. M. D.
Lesley A. Stevens, M. D., M. S.
Christopher H. Schmid, Ph. D.

Accuracy of a Glomerular Filtration Rate Estimating Equation over Time in People with a Wide Range of Kidney Function

Introduction: The change in glomerular filtration rate (GFR) is important for clinical decision making. However, the accuracy of GFR estimated from serum creatinine over time is not well known. The difference between measured GFR (mGFR) and estimated GFR (eGFR) (error) is usually attributed to non-GFR determinants of serum creatinine. We hypothesized that the mean error in a population would remain stable over time, but the inter-individual variation in the change over time in error would be large and related to clinical and demographic factors associated with non-GFR determinants of serum creatinine.

Methods: This is a longitudinal study of diagnostic accuracy including subjects from four studies with a wide range of kidney function. GFR was measured using urinary clearance of ^{125}I-iothalamate (reference test). GFR was estimated using the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation (index test). The change in error over time was modeled using longitudinal mixed models. Baseline covariates hypothesized to be associated with the non-GFR determinants of serum creatinine were tested in the mixed model.

Results: There were 13,708 GFR measurements in 3635 subjects over a mean follow up period of 3.6 years. In the pooled dataset the mean measured and estimated GFR and

error at baseline were 76, 76, and -0.3 ml/min/1.73 m^2. The mean change (standard error) in measured and estimated GFR and error were -2.3 (0.12), -2.2 (0.09) and -0.1 (0.10) ml/min/1.73 m^2 per year (P <.0001, <.0001, and 0.6 respectively). The variability (SD) among subjects in changes in measured and estimated GFR and error was 2.24, 1.59, and 1.91 ml/min/1.73 m^2 per year, respectively. Only 16% of subjects had changes in error larger than $\pm$ 3 ml/min/1.73 m^2 per year. A total of 8 non-GFR determinants were significantly associated with inter-individual variation in change in error in at least one study. Of the 8, only 1 explained greater than 20% of the variation [urine protein (22%)].

Conclusion: The accuracy of GFR estimates did not change over time in the population. Clinicians should interpret changes in estimated GFR over time as reflecting changes in measured GFR in most individuals.

Acknowledgements

I would like to acknowledge Drs. Lesley Stevens, Andrew Levey and Christopher Schmid for their extraordinary mentorship. The insight and support provided by them cannot be described in words. This project would have been unconceivable without their support. I am also grateful to Hocine Tighiouart and Jennifer Bassett Midle for all their help with SAS and Lucy Zhang for giving me the dataset. I would like to acknowledge Dr. David Kent who helped me shape this project and helped me stay focused on the main theme of the topic. I would like to express my gratitude to my fellow Sackler students especially Issa Dahabreh and my colleagues from the Division of Nephrology – Drs. Gautham Viswanathan, Navdeep Tangri and Shani Shastri for their support and precious input. I would like to acknowledge Nina Bonnoyer, Clinical Research Education Coordinator for Tufts CTSI for all her help whenever I needed her with Sackler school issues. Last but not the least; I am forever grateful to my family for understanding and always being there for me.

Table of Contents

List of Tables:

List of Figures:

List of Appendices:

Accuracy of a Glomerular Filtration Rate Estimating Equation over Time in People with a Wide Range of Kidney Function

INTRODUCTION

Evaluation and management of chronic kidney disease (CKD) requires monitoring of kidney function over time. The best overall measure of kidney function is the glomerular filtration rate (GFR) (Appendix 1). The gold standard for the measurement of GFR is the urinary clearance of an exogenous filtration marker, which is expensive, inconvenient, and may vary during the day. In clinical practice, serum levels of endogenous filtration markers, such as creatinine, are used to estimate the GFR. These serum levels are indirect measures of GFR as they are also affected by physiological processes other than GFR. For creatinine, the non-GFR determinants include its generation from dietary intake or muscle catabolism, tubular secretion, and extra-renal elimination (Appendices 2-3) [1-3]. Estimating equations use easily measured clinical variables as surrogates for these unmeasured non-GFR determinants, and provide more accurate estimates than the serum level alone [4-5]. These equations are widely used in clinical practice. Indeed serum creatinine is measured more than 280 million times per year in the US and estimated GFR (eGFR) based on serum creatinine is reported by more than 75% of clinical laboratories when serum creatinine is ordered [6-7].

While GFR estimating equations improve assessment of kidney function, the non-GFR determinants of serum creatinine may vary among individuals, leading to differences between measured GFR and estimated GFR (error). The non-GFR determinants also can vary over time within an individual, leading to changes in error over time. GFR estimating equations have been developed and extensively validated at single time points, but their performance over time is not well known.

In this study, we evaluate the accuracy of eGFR over time in diverse study populations with and without chronic kidney disease and diabetes over a wide range of measured GFR. We estimate GFR using the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation, based on age, sex, race and serum creatinine [8]. We hypothesized that the average error in the study population would remain stable over time but there would be wide variation within individuals in the change in the error over time, and that changes in error would be related in part to available clinical and demographic factors hypothesized to be non-GFR determinants of serum creatinine that are not included in the equation.

METHODS

Design

The design is a set of longitudinal studies of diagnostic accuracy using estimated GFR as the index test and measured GFR as the reference test. Our primary focus is the difference between measured and estimated GFR (error), changes in error, and factors associated with change in error over time. We considered a change in error smaller than $\pm$ 3 ml/min/1.73 m^2 year as clinically insignificant at all levels of measured GFR.

Study population

Studies included in the pooled data set are Modification of Diet in Renal Disease (MDRD) Study, African American Study of Kidney Disease and Hypertension (AASK), Collaborative Study Group (CSG), and Diabetes Complications and Control Trial

(DCCT). These studies, described in more detail in Appendices 4-5, were randomized, multicenter controlled trials and included patients with a wide range of GFR, diverse racial backgrounds and varied clinical characteristics. MDRD Study included individuals with chronic kidney disease. AASK was a study of hypertensive individuals, CSG was a study of type 1 diabetics with nephropathy and DCCT was a study of type 1 diabetics without nephropathy. In MDRD Study, data were available at 0, 6, 18 and 30 months. In AASK, data were available at 0, 3, 6, 12, 18, 24, 30, 36, 42, 48, 54 and 60 months. In CSG, data were available at 0, 12, 24 and 36 months. In DCCT, data were available at 0, 36 and 48 months.

Kidney Function Measures

GFR was measured using urinary clearance of ^{125}I-iothalamate and expressed adjusted for body surface area [9]. Serum creatinine (Scr) was assayed in the individual study laboratories and calibrated to standardized serum creatinine values ascertained by the Roche enzymatic method (Roche–Hitachi P-Module instrument with Roche Creatininase Plus assay, Hoffman-LaRoche, Basel, Switzerland) at the Cleveland Clinic Research Laboratory (Cleveland, Ohio) as described elsewhere [10]. Serum creatinine was expressed in mg/dl. Measures of GFR and serum creatinine were obtained longitudinally in each study.

GFR was estimated using the Chronic Kidney Disease Epidemiology (CKD-EPI) equation as $eGFR = 141 \times \min(Scr/\kappa, 1)^{\alpha} \times \max(Scr/\kappa, 1)^{-1.209} \times 0.993^{Age} \times 1.018$ [if female] × 1.159 [if black], where κ is 0.7 for females and 0.9 for males, α is -0.329 for females and -0.411 for males, min indicates the minimum of Scr/κ or 1; max indicates the

maximum of Scr/ κ or 1 8. As sex and race do not change over time, changes over time in estimated GFR reflect changes in age and serum creatinine.

Predictor variables

For evaluating the association between baseline demographic and clinical characteristics and errors in estimation of GFR, we selected candidate variables available in each study that were hypothesized to be associated with non-GFR determinants of serum creatinine based on a review of prior studies and clinical considerations [11]. Variables included are measures of muscle mass and body size (body surface area, body mass index, urine creatinine), measures related to severity of kidney disease (bicarbonate, phosphate, and urine protein), measures of inflammation (albumin, white blood cell count), and measures of dietary intake (urine phosphate, urine urea nitrogen) and other endogenous filtration markers (blood urea nitrogen).

Analytic plan

The primary outcome was error (difference between measured and estimated GFR). We summarized the descriptive data for measured GFR, estimated GFR and error using means and standard deviations for continuous data except when non-normality indicated medians and interquartile ranges would be more appropriate. In addition to mean error (bias), we also analyzed the median error, the mean of the absolute values of the error, interquartile range (IQR) of error and probability that error was within 30% of measured GFR (P30) [12-13]. Appendix 5 provides more details of the performance measures and their evaluation.

We explored the change over time of measured GFR, estimated GFR and error using mixed models that modeled each of these variables as a linear function of time for each study individually and for all studies pooled together. The mixed models permitted different random intercepts and slopes for each individual and allowed for correlation between these intercepts and slopes. We tested for heterogeneity among studies by incorporating an interaction between study and time and comparing models with and without the interaction using a likelihood ratio test. We performed complete case analysis to compare change in error over time among those who completed the study period versus those who did not. We also explored nonlinear patterns for change in error by testing quadratic terms in the pooled and individual studies.

We then incorporated the baseline characteristics as covariates in the mixed models for change in error over time. Covariates were centered on their study means. Urine protein was transformed to the logarithmic scale and centered on zero (i.e., on a urine protein of 1). We only assessed these associations in each individual study because of heterogeneity across studies in subject characteristics and lack of uniform availability of all variables across studies. We further explored if any of the non-GFR determinants explained non-linearity. Statistical analyses were performed in SAS version 9.2.

Further details on statistical methods can be found in Appendix 6.

RESULTS

Clinical Characteristics

The pooled data from the four studies included a total of 3635 subjects with measured GFR at baseline. Table 1 compares the clinical characteristics across the studies. The mean age was lower and the mean measured GFR was higher in CSG and DCCT compared to MDRD Study and AASK. The follow up time was 2.5 years in MDRD Study, 3 years in CSG, 4 years in DCCT and 5 years in AASK. Altogether there were 13708 GFR measurements in 3635 subjects during a mean follow up period of 3.5 years, including 2310 for MDRD Study, 8234 for AASK, 1138 for CSG and 2026 for DCCT (Table 2).

Changes in Measured and Estimated GFR and Error over Time

Descriptive data for measured GFR, estimated GFR and equation performance over time in each study (details in Appendix 5) show that both measured and estimated GFR decreased over time in MDRD, AASK and CSG. Measured GFR was stable over time in DCCT, whereas estimated GFR decreased. In each study, the performance of the CKD-EPI equation was consistent over time.

Table 3 and Figure 1 show results from the mixed model in the pooled dataset and by study. The mean measured and estimated GFR at baseline were both 76 ml/min/1.73 m^2, with mean (SE) rates of change in measured and estimated GFR of -2.3 (0.1), -2.2 (0.1) ml/min/1.73 m^2 per year, respectively. Mean error at baseline was very small [-0.3 (SE 0.27) ml/min/1.73 m^2] and did not change significantly over time [mean change in error of -0.1 (SE 0.10) ml/min/1.73 m^2] (Table 3). The variability (SD) among subjects in changes in measured GFR, estimated GFR and error was 2.24, 1.59 and 1.91 ml/min/1.73 m^2 per year, respectively. Only 16% of subjects had a change in error larger than $\pm$ 3

ml/min/1.73 m^2 per year (Figure 1). Appendix 6.1 shows the detailed results for individual variability.

There was substantial heterogeneity among studies ($p < 0.0001$; details found in Appendix 6.2). At baseline, mean error ranged from 4.9 ml/min per 1.73 m^2 for DCCT to -1.6 ml/min per 1.73 m^2 for CSG. For each study, the mean (SE) change in error per year was small [-0.9 (0.4) to 0.2 (0.1)] (Table 3). However, some individuals changed more than others (SD of the slope for MDRD, AASK, CSG and DCCT were 0.46, 0.78, 1.19 and 1.21 ml/min/1.73 m^2 per year respectively) (Figure 1, panel b, Appendix 6.1).

Association of Baseline Variables with Changes in Error over Time

Because of significant heterogeneity in change in error, analyses examining associations of predictor variables with change in error were performed in the individual studies. A total of eight baseline variables were found to be significantly associated with the variation between individual differences in change in error in at least one study (Table 4, and Appendix 6.3). In the MDRD Study, blood urea nitrogen, systolic blood pressure and urine protein were significant, but none of these accounted for more than 9% of inter-individual variation in the change in the error. In AASK, blood urea nitrogen and females were significant but explained only up to 5% of inter-individual variation in the change in the error. In CSG, body mass index, urine protein, serum albumin, blood urea nitrogen and serum phosphorus explained 15% to 22 % of inter-individual variation in the change in the error. In DCCT, systolic blood pressure and blood glucose were significant; but explained only 7% to 19% of inter-individual variation in change in error.

We assessed for non-linearity and found no effects that were clinically significant (Appendix 6.4). As a sensitivity analysis, we analyzed separately individuals with complete follow up. Results were consistent with our main analysis (Table 3, Appendix 6.5).

DISCUSSION

The mean error in GFR estimates was small and did not change substantially over time. Only 16% of subjects had a change in error that was larger than $\pm$ 3 ml/min/1.73 m^2, a magnitude that we considered clinical significant. The change in error was not consistently related to any particular variable determined at baseline. Thus, in these studies, the CKD-EPI equation provided unbiased but imprecise estimates of measured GFR over time. These findings are consistent with the previous literature on the accuracy of the CKD-EPI equation at a single time point [8]. These findings have important implications for interpretation of GFR estimates in clinical practice.

In principle, error in GFR estimates reflects the effect of non-GFR determinants of creatinine unaccounted for by variables in the GFR estimating equation as well as errors in measured GFR or serum creatinine. To avoid systematic errors in measured GFR or serum creatinine, we used a single exogenous filtration marker across all studies and at all time points and calibrated creatinine assays in each study to reference materials at baseline and applied that calibration to all future time points. It is therefore most likely that the imprecision we observed in estimates of measured GFR over time reflect random changes in non-GFR determinants or in GFR measurement error.

There are limited data on variation in non-GFR determinants of serum creatinine over time. Studies of variability over time in urinary creatinine excretion in normal subjects have shown intra-individual coefficients of variation range from 10.5-14.4%, suggesting random variation in creatinine generation [14-17]. There are fewer data on variation in creatinine secretion. Systematic variation in dietary intake can affect both creatinine generation and secretion [18]. Thus, it is to be expected that patient characteristics associated with changes in non-GFR determinants of serum creatinine would be associated with changes in error over time. None of the subject characteristics at baseline that we examined explained why some individuals changed more than others suggesting that variation in non-GFR determinants may not be the cause of variation in change in error over time. However, our analysis is limited, because we did not explore changes in these subject characteristics and their association to error over time as we did not have uniform follow up data on these variables. It is also possible that the observed changes are due to regression to the mean in an individual patient.

Previous studies suggest that there is a reasonable amount of variability in measured GFR, due either to biological variability or measurement error. In a classic study using urinary clearance of inulin, Homer Smith documented a coefficient of variation (standard deviation divided by the mean) for repeated GFR measurements in a single individual over time to be approximately 7.5% [19]. Other reports from the MDRD Study also show substantial variability in measured GFR over time. One study found that between-day coefficients of variations of ^{125}I-iothalamate clearance were within the range of 11.6% and 16.6% [20]. One study reported a median coefficient of variation between 2 GFR measures 3 months apart of 6.3 %, and a more recent study using data from MDRD Study

and AASK combined showed a coefficient of variation of 11.9% of measures 2 months apart [18, 21-22]. Therefore, it is likely that a substantial portion of the variation in change in error over time that we observed reflects GFR measurement error, rather than changes in non-GFR determinants in serum creatinine.

Previous studies have evaluated the performance of GFR estimating equations over time in the MDRD Study and in AASK using other statistical methods. In a previous report from the MDRD Study, rates of decline in measured vs. estimated GFR were compared using marginal models [23]. The authors showed a mean (SD) decline in measured and estimated GFR of 3.9 (7.2) vs. 2.8 (7.1) ml/min per 1.73 m^2 per year, which are larger than the values reported here. The mixed model that we used is more robust for assessing changes over time in subjects with varying amount of follow-up, as observed in our study population. As reported in our study, these authors also found that differences in slope estimates between measured and estimated GFR was not related to a large number of baseline factors. In another study in AASK, time-to-event outcomes were compared using rates of changes in measured and estimated GFR [24]. The association of baseline factors were similar with events defined by estimated GFR and measured GFR (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. Overall, both of these reports are consistent with our findings and suggest that baseline factors do not appear to contribute substantially to differences in estimated vs. measured GFR over time.

GFR estimates are used in most clinical circumstances each day. For example, they are used to make decisions regarding detection of CKD or acute kidney injury, use of iodinated or gadolinium contrast agents; and dosing of medications. Less frequently but importantly, they are also used to determine optimal timing of pre-emptive transplant or listing for a cadaveric kidney transplant, and initiation of dialysis. Changes in GFR estimates are central to these decisions. It is thus important that a change in estimated GFR reflects a change in measured GFR. Although as we demonstrated, GFR estimates are on average, unbiased over time, with only 16% of changes in error large enough to be clinically significant. Moreover, changes in error could not be accounted for by patients clinical characteristics at baseline. As such, for routine clinical decision making, clinicians should interpret a change in estimated GFR as a reflection of a change in measured GFR, and act accordingly. However, in patients in whom clinical circumstances suggest a change in non-GFR determinants of serum creatinine, for example, recent hospitalizations, decreased oral intake or refeeding after an illness, muscle mass loss as with illness or amputation, then the change in estimated GFR could reflect the change in non-GFR determinants of serum creatinine rather than a change in measured GFR [25]. In such patients, clinicians should consider measuring GFR as a confirmatory test if more accurate information would improve clinical decision making.

The strengths of our study include large pooled data set which allowed analysis of a very large number of GFR measurements, in a multiethnic setting, with moderately and severely reduced kidney function; a wide range of kidney disease diagnoses including diabetics who contribute a major share to the CKD burden; and sequential data from longitudinal follow up. These studies have a similar protocol for measuring GFR and

calibration of serum creatinine which allowed for uniformity of analysis and interpretation. The use of CKD-EPI equation in the same study cohorts in which it was developed is a strength of our study, as it allows for a good fit at baseline, facilitating identification of deviations in fit over time. Finally, the use of a mixed model is a robust technique to evaluate both within and between individual variations in the change in error.

We acknowledge the following limitations in our study: First, the covariates were not available across all data sets, requiring analyses stratified by study. Also, we could not assess changes in subject characteristics over time. Second, although we had a diverse individual representation in terms of racial backgrounds and underlying disease, we had limited information on Hispanics, Asians and transplant recipients. Third, we did not have repeated measurements over short intervals to evaluate biologic variability in measured GFR and random error in GFR measurement.

In summary, the CKD-EPI equation performs well over time. The finding of unbiased GFR estimates over time on average suggests that changes in estimated GFR reflect changes in measured GFR rather than changes in non-GFR determinants of serum creatinine.

Table 1: Subject Characteristics at baseline

Study Name	**MDRD Study**		**AASK**		**CSG**		**DCCT**	
	Mean (SD) range or %	*N*	*Mean (SD) range or %*	*N*	*Mean (SD) range or %*	*N*	*Mean (SD) range or %*	*N*
Number of individuals	NA	831	NA	1029	NA	401	NA	1381
Age range (years)	18-70	142[+]	18-70	202[+]	18-49	NA	13-39	NA
Sex (%male)	60	503	62	636	54	216	69	740
Race (%)	80% White	766	100% Black	1029	89% White	369	86% White	1333
Measured GFR (ml/min/1.73m^2)	33 (12)	831	46 (13)	1029	77 (33)	401	125 (22)	1378
Estimated GFR (ml/min/1.73m^2)	36 (14)	831	15 (16)	1029	79 (27)	401	120 (14)	1377
Serum Creatinine (mg/dl)	2.3 (0.9)	831	1.9 (0.7)	1029	1.3 (0.4)	401	0.8 (0.1)	1377
Body Mass Index (kg/m^2)	27 (4.5)	829	NA	NA	25 (4.1)	400	24 (3.0)	1380
Body Surface Area (m^2)	1.9 (0.2)	831	2.0 (0.2)	1029	1.8 (0.2)	401	1.8 (0.2)	1378
Systolic Blood Pressure (mmHg)	134 (19.6)	829	150 (24.0)	1029	138 (19.5)	401	114 (11.6)	1361
Diastolic Blood Pressure (mmHg)	NA	NA	95 (14.3)	1029	86 (11.2)	401	73 (8.4)	1362
Blood Urea Nitrogen (mg/dl)	34 (13.0)	831	23 (10.0)	1029	23 (11.6)	400	NA	NA
Serum Glucose (mg/dl)	NA	NA	NA	NA	NA	NA	200 (81.5)	937
Bicarbonate (mEq/L)	23 (3.7)	820	25 (3.0)	1029	27 (3.8)	393	NA	NA
Albumin (g/dl)	4.0 (0.4)	831	4.2 (0.4)	1029	3.7 (0.5)	397	3.9 (0.3)	1377
Phosphate (mg/dl)	3.7 (0.7)	826	3.5 (0.6)	1029	3.7 (0.7)	397	NA	NA
Total Cholesterol (mg/dl)	220 (48.6)	826	212 (44.2)	1020	236 (67.5)	399	180 (34.9)	1370
Hemoglobin (g/dl)	13.1(1.8)	800	NA	NA	13.4 (2.0)	397	NA	NA
White Blood Cell Count (K/mm^3)	6.7 (2.0)	810	NA	NA	7.7 (2.2)	400	NA	NA
Urine Volume (ml/d)	2679 (903)	806	2362 (950)	429	2448 (1026)	400	2340 (981)	503
Urine Creatinine (mg/d)	1407 (413)	806	1644 (616)	429	1607 (2030)	392	1444 (493)	533
Urine Protein (g/d)	1.1 (1.7)	806	0.5 (0.9)	1029	2.9 (2.9)	399	0.0 (0.1)	1376
Urine Phosphate (mg/d)	822 (285)	806	NA	NA	919 (1375)	87	NA	NA
Urine Urea Nitrogen (g/d)	9.2 (2.9)	806	8.4 (3.8)	429	11.6 (10.8)	124	10.2 (3.8)	39

[+]Older individuals > 65 years

MDRD - Modification of Diet in Renal Disease; AASK - African American Study of Kidney Disease and Hypertension; CSG - Collaborative Study Group; DCCT - Diabetes Complications and Control Trial; NA – Not available; SD - Standard Deviation

Table 2: Number of Subjects at each Time by Study

Time	MDRD Study	AASK	CSG	DCCT
Baseline	831	1029	401	1381
3 months	NA	838	NA	NA
6 months	665	861	NA	NA
12 months	NA	859	310	NA
18 months	523	787	NA	NA
24 months	NA	736	272	NA
30 months	291	697	NA	NA
36 months	NA	686	155	501
42 months	NA	561	NA	NA
48 months	NA	473	NA	151
54 months	NA	392	NA	NA
60 months	NA	315	NA	NA

MDRD - Modification of Diet in Renal Disease; AASK - African American Study of Kidney Disease and Hypertension; CSG - Collaborative Study Group; DCCT - Diabetes Complications and Control Trial; NA – Not available.

Table 3: Measured and Estimated GFR and Difference (Error) at the First Visit and Rate of Change over Time, Overall and by Study

		Population (N)				
Kidney Function Measures		**Pooled (3635)** *Coefficient (SE)*	**MDRD Study (831)** *Coefficient (SE)*	**AASK (1029)** *Coefficient (SE)*	**CSG (401)** *Coefficient (SE)*	**DCCT (1374)** *Coefficient (SE)*
Measured GFR	Mean*	**76 (0.7)**	**33 (0.4)**	**46 (0.4)**	**76 (1.6)**	**125 (0.6)**
	Change+	**-2.3 (0.1)**	**-3.3 (0.2)**	**-1.9 (0.1)**	**-5.7 (0.5)**	**-0.9 (0.3)**
Estimated GFR	Mean	**76 (0.7)**	**35 (0.5)**	**50 (0.5)**	**78 (1.3)**	**120 (0.4)**
	Change+	**-2.2 (0.1)**	**-3.5 (0.2)**	**-2.2 (0.1)**	**-4.8 (0.4)**	**-1.5 (0.1)**
Error	Mean	-0.3(0.3)	**-2.4 (0.2)**	**-4.1 (0.3)**	-1.6 (0.9)	**4.9 (0.6)**
	Change+	-0.1 (0.1)	0.2 (0.1)	**0.3 (0.1)**	**-0.9 (0.4)**	**0.6 (0.3)**

* Unit for GFR and error is ml/min/1.73 m^2

+Unit for change is ml/min/1.73 m^2 per year

-Bold indicates p-value < 0.05 for the coefficient

MDRD - Modification of Diet in Renal Disease, AASK - African American Study of Kidney Disease and Hypertension, CSG - Collaborative Study Group, DCCT - Diabetes Complications and Control Trial

Table 4: Proportion of Variability in the Change in Error Attributable to Variation in Baseline Characteristics

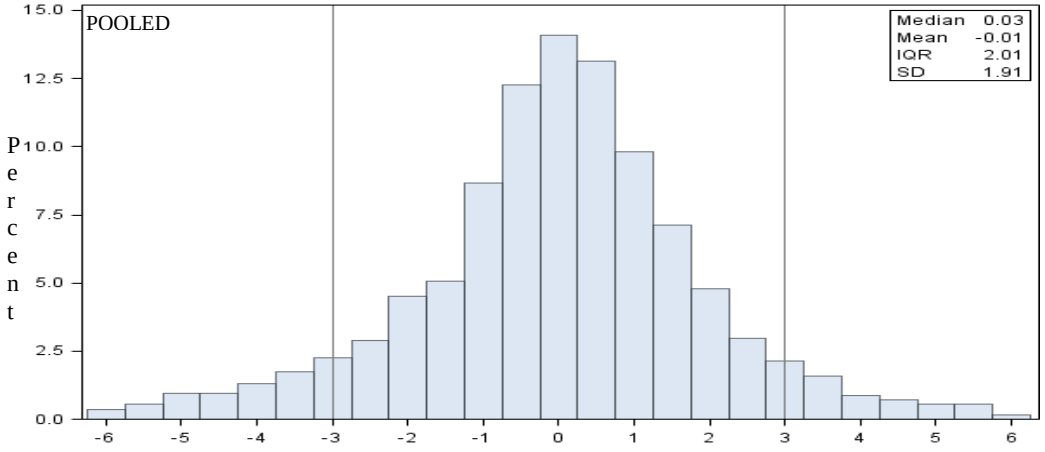
Variables	**MDRD Study**	**AASK**	**CSG**	**DCCT**
Females	2%	**4%**	0.4%	-1%
Systolic Blood Pressure (mmHg)	**4%**	1%	-0.5%	**7%**
Body Mass Index (kg/m^2)	2%	NA	**20%**	2%
Blood Urea Nitrogen (mg/dl)	**5%**	**5%**	**15%**	NA
Albumin (g/dl)	-0.4%	0.1%	**16%**	1%
Phosphate (mg/dl)	2%	0.5%	**16%**	NA
Serum Glucose (mg/dl)	NA	NA	NA	**19%**
Urine Protein (g/d)	**9%**	1%	**22%**	-2%

-The numbers explain the percent change in variance component explained by the covariates.
-Variables with significant interaction with time are in bold
MDRD - Modification of Diet in Renal Disease, AASK - African American Study of Kidney Disease and Hypertension, CSG - Collaborative Study Group, DCCT - Diabetes Complications and Control Trial
NA -Covariate information not available in the dataset

Figure 1: Distribution of Change in Error - Pooled and by Study

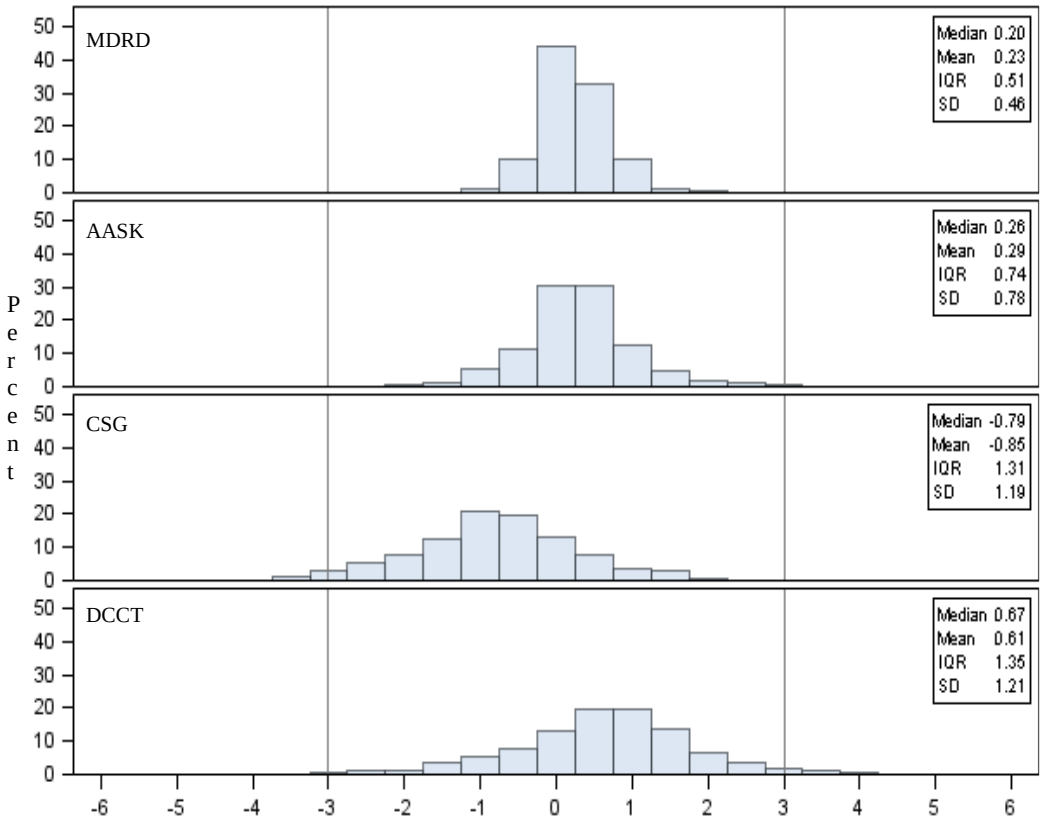
The bars are the random patient errors. The x-axis is truncated at ± 6 ml/min/1.73 m^2 per year

Figure 1a: Pooled -1.02 (25th percentile), 1.11 (75th percentile) ml/min/1.73 m^2 per year.



Within Individual Change in Error per Year

Figure 1b: By study. MDRD Study (25th percentile, 75th percentile): -0.01, 0.46 ml/min/1.73 m^2 per year, AASK (25th percentile, 75th percentile): -0.09, 0.65 ml/min/1.73 m^2 per year, CSG (25th percentile, 75th percentile): -1.44, -0.12 ml/min/1.73 m^2 per year, DCCT (25th percentile, 75th percentile): 0.00, 1.35 ml/min/1.73 m^2 per year



Within Individual Change in Error per Year

Appendix 1: Importance of GFR in Clinical Decisions

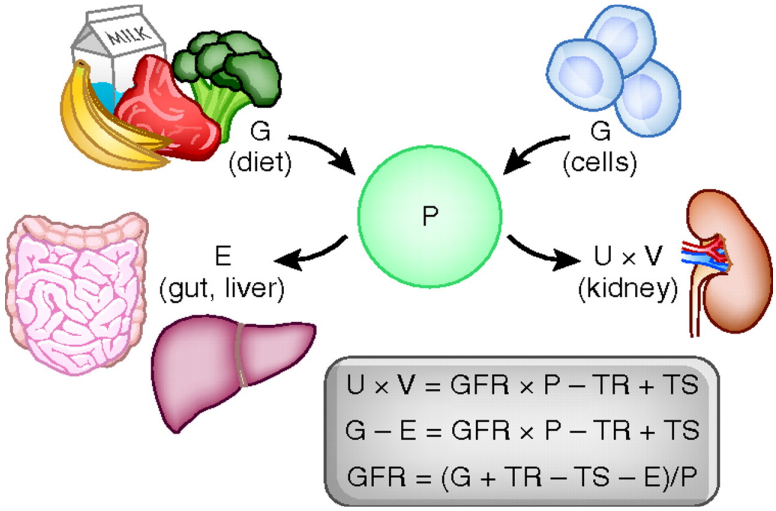
GFR is considered the best overall index of kidney function. GFR is used not only in the detection but also in the prognosis and management of CKD. The table below shows various clinical situations where the use of GFR is relevant.

Clinical Conditions where Assessment of GFR is Important [25]

Clinical Decisions	Current Level of GFR	Change in Level of GFR
Diagnosis	• Detection of CKD • Evaluation for kidney donation	• Detection of AKI • Detection of CKD progression
Prognosis	• Risk of CKD complications • Risk for CVD • Risk for mortality	• Risk for kidney failure
Treatment	• Dosage and monitoring for medications cleared by the kidney • Determine safety of diagnostic tests or procedures • Referral to nephrologists • Referral for kidney transplantation • Placement of dialysis access	• Treatment of AKI • Monitoring drug toxicity

Appendix 2: Determinants of Serum Levels of Endogenous Filtration Markers

The plasma level (P) of an endogenous filtration marker is determined by its generation (G) from cells and diet, extra-renal elimination (E) by gut and liver, and urinary excretion (UV) by the kidney. Urinary excretion is the sum of filtered load (GFR X P), tubular secretion (TS) and reabsorption (TR). In the steady state, urinary excretion equals generation and extra-renal elimination. By substitution and re-arrangement, GFR can be expressed as the ratio of the non-GFR determinants (G, TS, TR and E) to the plasma level [25].



Appendix 3: Non-GFR Determinants of Serum Creatinine [25]

Creatinine-based estimating equations include age, sex, race or weight as surrogates for differences in creatinine generation from muscle mass [5, 26]. However, there are other non-GFR determinants of creatinine as shown in the table below that have not been accounted for in the GFR estimating equations and could potentially lead to errors in estimated GFR.

Factors[+]	**Effect on serum creatinine independent of GFR**		**Accounted for in GFR estimating equations**
	Direction	**Mechanism**	
Age	Decrease	Generation	Yes
Female Sex	Decrease	Generation	Yes
Race			
African American	Increase	Generation	Yes
Hispanics	Decrease		No
Asian	Increase/ Decrease		Yes
Body Habitus			
Muscular	Increase	Generation	No
Amputation	Decrease		No
Obesity	No change		No
Chronic Illness			
Malnutrition, inflammation, de-conditioning	Decrease	Generation	No
Neuromuscular diseases	Decrease		No
Liver disease	Decrease		No
HIV	Decrease		No
Diet			
Vegetarian Diet	Decrease	Generation	No
Ingestion of Cooked Meat	Increase		No
Medications			
Cimetidine	Increase	Tubular secretion	No
Trimethroprim	Increase	Tubular secretion	No
Antibiotics	Increase	Extra-renal elimination	No

+Factors are the non-GFR determinants of Serum Creatinine

Appendix 4: Sources of Data

We identified studies from the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) with measurement of GFR using urinary clearance of ^{125}I-iothalamate, ability to calibrate serum creatinine, wide GFR range, kidney pathology and race. The details on randomization, the design and baseline characteristics have been previously published.

Modification of Diet in Renal Disease Study (MDRD Study)

MDRD study was a randomized, multicenter controlled trial of protein intake and blood pressure control of patients with moderate (study A) to severe (study B) non-diabetic kidney disease with a 2x2 factorial design. The study was conducted between 1989 and 1994. The participants were stratified to study A with initial mGFR of 25-55 ml/min/1.73 m^2 and study B with initial mGFR of 13-24 ml/min/1.73 m^2. The age range was 18-70 years with 80% Caucasians and diverse kidney pathology. The subjects had a follow-up of 4 years. For our analysis we had 2.5 years of data available to us [27-28].

African American Study of Kidney Disease and Hypertension (AASK)

AASK was a randomized, multicenter controlled trial in black individuals of effectiveness of three antihypertensive regimens (ramipril, amlodipine and metoprolol) and two levels of BP control (mean arterial pressure of < 92 versus 102 to 107 mmHg) on the progression of hypertensive kidney disease with a 3x2 factorial design. The study was conducted from 1994 to 2001. The mean measured GFR was 46 ml/min/1.73 m^2, the age range was 18-70 years and the presumed diagnosis for the underlying kidney pathology was hypertensive nephrosclerosis. The subjects had a follow-up of 6 years; however, we

used 5 years of follow-up data for our analysis. The GFR was measured at baseline, at 3 and 6 months, then every 6 months thereafter [29-30].

Collaborative Study Group (CSG)

CSG was a randomized double blind placebo controlled multicenter study of type -1 diabetics comparing captopril with placebo to determine whether captopril has kidney-protecting properties independent of its effect on blood pressure in diabetic nephropathy. The mean measured GFR was 77 ml/min/1.73 m^2. The age range was 18-49 years and subjects were predominantly Caucasians. The median follow-up was 3 years with yearly GFR measurements from the first year onwards until 3 years [31].

Diabetes Complications and Control Trial (DCCT)

DCCT was a randomized controlled multicenter trial to study the effects of intensive insulin therapy on the development and progression of microvascular complications in type 1 diabetics. The subjects were randomized to standard therapy or intensive control therapy. The study was conducted from 1983 to 1993. The trial involved 1,441 volunteers, ages 13 to 39, with mean measured GFR of 125 ml/min/1.73 m^2 and included 86 % Caucasians. The subjects had diabetes for at least 1 year but no longer than 15 years and negligible to minimal proteinuria. GFR was measured at study entry, at years 3 and 4 - but was not measured in all subjects [32-34].

Appendix 5: Performance Metrics of Estimated GFR at Each Time Point by Study

We summarized the cross-sectional relationship between measured and estimated GFR among all participants available at each time point using the performance metrics described in the methods section. These fall into three main categories: bias, precision and accuracy. Bias, although technically defined as the mean difference between measured and estimated GFR, as a general concept describes the central tendency of the errors, whether expressed as a mean, median or mean of absolute values. Bias arises when systematic errors cause the estimation to consistently miss its targeted measure. Precision describes the variability of the differences about the average difference, expressed either as a standard deviation or an interquartile range. Variability arises from imprecise measurements, measurement errors or inaccurate models to estimate GFR. Accuracy summarized using P30 which is the percentage of estimated GFR within 30% of measured GFR, incorporates bias and precision and therefore reflects both systematic and random error. P30 is a quantile based measure and is robust to outliers [12].

MDRD Study

Over 30 months of follow up, serum creatinine increased, measured and estimated GFR decreased concomitantly. The mean bias was -2.2 ml/min/1.73 m^2 at baseline and at 30 months was -1.9 ml/min/1.73 m^2. The IQR was 7.5 ml/min/1.73 m^2 at baseline and was 6.7 ml/min/1.73 m^2 at 30 months. Accuracy started at 87% at baseline but declined to 81% at the end of 30 months (Appendix 5, Table a).

Appendix 5, Table a

	Month 0	Month 6	Month 18	Month 30
Sample Size	831	665	523	291
Mean Scr (mg/dl)	2.3	2.5	2.8	2.9
Mean mGFR (ml/min/1.73 m^2)	34	31	28	28
Mean eGFR (ml/min/1.73 m^2)	36	33	30	30
Mean Error (ml/min/1.73 m^2)	-2.2	-2.7	-2.0	-1.9
Median Error (ml/min/1.73 m^2)	-1.6	-2.1	-1.3	-1.0
Absolute bias (ml/min/1.73 m^2)	5.1	4.9	4.3	4.8
IQR (ml/min/1.73 m^2) of Error	7.5	6.9	6.0	6.7
SD (ml/min/1.73 m^2) of Error	6.7	6.3	5.7	6.5
P30 (%)	87	83	84	81

Scr-Serum Creatinine, mGFR-measured GFR, eGFR-estimated GFR, Absolute Bias was the absolute mean bias, IQR – Interquartile Range, SD – Standard Deviation, P30 – Percentage of eGFR within 30% of mGFR

AASK

Over 60 months of follow up, serum creatinine increased and measured and estimated GFR decreased slightly. The mean bias was -4.2 ml/min/1.73 m^2 at baseline and was -2.5 ml/min/1.73 m^2 at 60 months. The IQR was 12.2 ml/min/1.73 m^2 and accuracy was 83% at baseline and at 60 months was 11.0 ml/min/1.73 m^2 and 78%, respectively.

Appendix 5, Table b

	Month 0	Month 3	Month 6	Month 12	Month 18	Month 24	Month 30	Month 36	Month 42	Month 48	Month 54	Month 60
Sample Size	1029	838	861	859	787	736	697	686	561	473	392	315
Mean Scr (mg/dl)	2.0	2.0	2.1	2.1	2.2	2.2	2.2	2.2	2.1	2.2	2.2	2.3
Mean mGFR (ml/min/1.73 m^2)	46	46	46	46	45	45	45	44	45	45	45	43
Mean eGFR (ml/min/1.73 m^2)	50	50	49	49	49	49	49	48	49	48	47	46
Mean Error (ml/min/1.73 m^2)	-4.2	-3.7	-3.5	-3.6	-3.6	-3.9	-4.1	-3.9	-3.4	-3.2	-2.5	-2.5
Median Error (ml/min/1.73 m^2)	-3.0	-3.0	-3.0	-3.0	-2.9	-2.7	-2.9	-3.2	-2.2	-2.3	-1.5	-2.6
Absolute bias (ml/min/1.73 m^2)	8.3	7.9	8.2	8.1	7.4	7.9	8.2	8.0	8.1	8.1	7.3	8.1
IQR (ml/min/1.73 m^2) of Error	12.2	11.1	12.2	11.74	10.8	10.5	11.5	11.6	11.4	10.9	10.4	11.0
SD (ml/min/1.73 m^2) of Error	10.5	10.0	10.4	10.5	9.3	10.4	10.8	10.3	11.5	11.3	10.1	11.1
P30 (%)	83	80	82	81	82	81	79	79	81	81	82	78

Scr-Serum Creatinine, mGFR-measured GFR, eGFR-estimated GFR, Absolute Bias was the absolute mean bias, IQR – Interquartile Range, SD – Standard Deviation, P30 –Percentage of eGFR within 30% of mGFR

CSG

36 months of follow up, serum creatinine on an average increased and measured and estimated GFR decreased concomitantly. The mean bias was -2 ml/min/1.73 m^2 at baseline and was -3.5 ml/min/1.73 m^2 at 36 months. Accuracy was 81% at baseline and at 36 months declined to 77%.

Appendix 5, Table c

	Month 0	Month 12	Month 24	Month 36
Sample Size	401	310	272	155
Mean Scr (mg/dl)	1.3	1.4	1.5	1.5
Mean mGFR (ml/min/1.73 m^2)	77	71	68	67
Mean eGFR (ml/min/1.73 m^2)	79	72	72	71
Mean Error (ml/min/1.73 m^2)	-2.0	-0.7	-3.3	-3.5
Median Error (ml/min/1.73 m^2)	-3.2	-1.7	-3.5	-4.3
Absolute bias (ml/min/1.73 m^2)	13.6	12.9	11.8	11.9
IQR (ml/min/1.73 m^2) of Error	19.7	17.0	15.4	18.4
SD (ml/min/1.73 m^2) of Error	19.0	19.3	16.5	15.4
P30 (%)	81	81	76	77

Scr-Serum Creatinine, mGFR-measured GFR, eGFR-estimated GFR, Absolute Bias was the absolute mean bias, IQR – Interquartile Range, SD – Standard Deviation, P30 – Percentage of eGFR within 30% of mGFR

DCCT

Over a follow up of 48 months, average serum creatinine and measured and estimated GFR remain stable. The mean bias was 2.6 ml/min/1.73 m^2 at baseline and at 48 months was 5.2 ml/min/1.73 m^2. Accuracy remained unchanged at each time point.

Appendix 5, Table d

	Month 0	Month 36	Month 48
Sample Size	1374	501	151
Mean Scr (mg/dl)	0.8	0.9	0.9

Mean mGFR (ml/min/1.73 m^2)	126	123	126
Mean eGFR (ml/min/1.73 m^2)	123	123	121
Mean Error (ml/min/1.73 m^2)	2.6	-0.2	5.2
Median Error (ml/min/1.73 m^2)	1.6	-1.3	2.2
Absolute bias (ml/min/1.73 m^2)	17.3	16.0	16.7
IQR (ml/min/1.73 m^2) of Error	26.4	25.2	27.5
SD (ml/min/1.73 m^2) of Error	23.0	21.2	23.9
P30 (%)	91	90	91

Scr-Serum Creatinine, mGFR-measured GFR, eGFR-estimated GFR, Absolute Bias was the absolute mean bias, IQR – Interquartile Range, SD – Standard Deviation, P30 – Percentage of eGFR within 30% of mGFR

Appendix 6: Detailed Statistical Methods

To fit the mixed models, we created a person-period data set in which each individual had one record for every time point. We fit a series of models to estimate changes over time in measured GFR, estimated GFR and error for each individual and the average across all individuals, overall and by studies [35]. Because of the small number of measurements on many individuals, we focused on linear and quadratic relationships

6.1 Within Individual Variability in Change in Error Explained by Time

To establish whether there was any systematic variation in the measured GFR, estimated GFR and mean error and to see if the variation resided within or between individuals and to explore whether there was any proportional reduction in variability in error with time, we fit a means model and a growth model. The means model assumes no relationship with time; the growth model assumes that an individual's mean error is a linear (or quadratic) function of time. For simplicity, we present the models for mean error below, but they apply to both measured and estimated GFR as well.

The means model for the mean error, Y_{it}, for the i^{th} individual at time t takes the form:

$$Y_{it} = \pi_{i0} + e_{it} \quad \text{(level 1)} \qquad \text{(A6.1.1)}$$

$$\pi_{i0} = \gamma_{00} + \zeta_{0i} \quad \text{(level 2)} \qquad \text{(A6.1.2)}$$

where $e_{it} \sim N(0, \sigma_e^2)$ and $\zeta_{0i} \sim N(0, \sigma_0^2)$

In this model, the true error for individual i is π_{i0} and the true error across all individuals is γ_{00}. On occasion t, the observed error Y_{it} deviates from the i^{th} individual's true error π_{i0} by the within-individual residual e_{it} which has mean 0 and variance σ_e^2 that describes the scatter of the individual time-specific errors around their own mean. For person i, the true

individual specific mean (π_{i0}) deviates from the population average true mean γ_{00} by the level-2 residual ζ_{0i} which has mean 0 and variance σ_0^2, the between scatter of individual-specific means around the population mean).

In the growth model, time is inserted as a predictor in level-1. For now, we include no substantive predictors at level-2, so comparison of the growth and means models evaluates how time can explain within-individual variation. The growth model has the form:

$$Y_{it} = \pi_{i0} + \pi_{i1}\text{time}_{it} + e_{it} \quad \text{(level 1)} \qquad \text{(A6.1.3)}$$

$$\pi_{i0} = \gamma_{00} + \zeta_{0i} \quad \text{(level 2)} \qquad \text{(A6.1.4)}$$

$$\pi_{i1} = \gamma_{10} + \zeta_{1i}. \qquad \text{(A6.1.5)}$$

In this model, ζ_{0i} is the between-individual error in the baseline error at time 0 and ζ_{1i} is the between-individual error in the change over time. We have

$$\begin{bmatrix} \zeta_{0i} \\ \zeta_{1i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{10} & \sigma_1^2 \end{bmatrix} \right) \qquad \text{(A6.1.6)}$$

so that there are variance components for between-individual intercepts and slopes as well as a correlation between the two. Now, σ_{e1}^2 is the within-individual residual variance that summarizes the scatter of the errors around the linear change trajectories.
We explore the proportional reduction in residual variance with addition of time using a pseudo-R^2 statistic, $(\sigma_e^2 - \sigma_{e1}^2) / \sigma_{e1}^2$.

In the table below, we show the within-individual residual variance from the first and second models and the percent reduction in the variance with addition of time in the second model.

	Pooled	MDRD Study	AASK	CSG	DCCT
Variation at baseline (σ_{e1}^2, means model)	81.77	18.89	55.35	185.16	355.49

Variation over time (σ^2_{e1}, growth model)	62.58	17.65	50.81	173.19	303.26
Percent reduction	23%	6.5%	8.2%	6.5%	14.7%

We conclude that 6.5% to 14.7% of the within-individual variation in error is explained by time in the individual studies and 23% of the within-individual variation in error is explained by time in the pooled data. The only way to reduce this within-individual variance further is to add time-varying covariates to the level-1 model. Since we have only baseline covariates, the within-individual residual variance remains unchanged in the models with the addition of baseline covariates.

6.2 Analysis Stratified by Study

Next, we tested to see if the individual studies comprising the pooled dataset were different. We used likelihood ratio tests. Our null hypothesis was that the studies were homogenous in the rate of change over time. We modeled the bias in the i^{th} individual at time t as a function of time as in equation A6.1.3 above.

Each individual's intercepts and slopes were modeled as a function of study as:

$$\pi_{i0} = \gamma_{00} + \gamma_{01}\text{MDRD}_i + \gamma_{02}\text{AASK}_i + \gamma_{03}\text{CSG}_i + \zeta_{0i} \qquad \text{(A6.2.1)}$$

$$\pi_{i1} = \gamma_{10} + \gamma_{11}\text{MDRD}_i + \gamma_{12}\text{AASK}_i + \gamma_{13}\text{CSG}_i + \zeta_{1i} \qquad \text{(A6.2.2)}$$

with covariance matrix

$$\begin{bmatrix} \zeta_{0i} \\ \zeta_{1i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{10} & \sigma_1^2 \end{bmatrix} \right)$$

The combined mixed model then has fixed effects of study and time as well as interactions of study with time and random effects for the intercept and time.

Fitting this model to each study, we have:

	Baseline Mean Error (γ_{00})*	Slope/ Rate of Change in Error (γ_{10})**
Pooled Model	-0.3 (0.3)	-0.1 (0.1)
Combined Model^{+}		
MDRD Study γ_{01},γ_{11}	**-2.8 (0.6)**	-0.1 (0.3)
AASK γ_{02}, γ_{12}	**-3.5 (0.6)**	-0.2 (0.2)
CSG γ_{03}, γ_{13}	**-2.5 (0.9)**	**-1.3 (0.4)**
DCCT γ_{00}, γ_{10}	**5.0 (0.4)**	**0.4 (0.2)**

* Unit for mean error is ml/min/1.73 m^2; **Unit for rate of change in error is ml/min/1.73 m^2 per year
-Coefficient with p-value <0.05 is in bold. +By the likelihood ratio test, the difference between the two models is 332.1 with 6 degrees of freedom, (p <0.0001)

Appendix 6.3: Associations of Baseline Characteristics with the Change in Error

The table below shows the association of baseline covariates with the random person-specific intercepts and slopes. The covariates are level-2 predictors and describe the between individual differences in the rate of change in error.. The model was of the form:

$$Y_{it} = \pi_{i0} + \pi_{i1*}\text{Time}_{it} + e_{it}$$

$$\pi_{i0} = \gamma_{00} + \gamma_{01}\text{Covariate}_i + \zeta_{0i}$$

$$\pi_{i1} = \gamma_{10} + \gamma_{11}\text{Covariate}_i + \zeta_{1i}$$

Of the 19 covariates that were tested, 8 were found to be significantly associated with time in at least one study Shown here are the fixed effects of the covariates and their interactions with time. The intercepts γ_{00}were calculated but are not shown in the table

below. γ_{01} is the estimate of the covariate, γ_{10} is the estimate of time and γ_{11} is the estimate of the interaction of time with the covariate.

Appendix 6.3, Table 1: Results from mixed models assessing effect of baseline covariates on error.

Covariate		**MDRD Study**	**AASK**	**CSG**	**DCCT**
		Coefficient (SE)	*Coefficient (SE)*	*Coefficient (SE)*	*Coefficient (SE)*
Body Mass Index (kg/m^2)	Covariate	**0.14 (0.05)**	NA	0.13 (0.22)	**0.52 (0.20)**
	Time*	0.23 (0.12)	NA	**-0.94 (0.41)**	0.62 (0.31)
	Interaction	-0.03 (0.03)	NA	**0.27 (0.11)**	-0.13 (0.11)
Body Surface Area (m^2)	Covariate	**3.01 (0.94)**	**4.74 (1.11)**	-0.57 (4.54)	-0.37 (3.09)
	Time*	0.23 (0.12)	**0.29 (0.08)**	**-0.90 (0.42)**	**0.61 (0.31)**
	Interaction	0.08 (0.52)	-0.27 (0.32)	2.74 (2.22)	-2.86 (1.56)
Systolic Blood Pressure (mmHg)	Covariate	**0.03 (0.01)**	0.01 (0.1)	**-0.13 (0.05)**	**0.13 (0.05)**
	Time*	0.22 (0.12)	**0.29 (0.08)**	**-0.87 (0.42)**	**0.63 (0.31)**
	Interaction	**-0.01 (0.01)**	-0.00 (0.00)	0.01 (0.02)	**-0.06 (0.03)**
Diastolic Blood Pressure (mmHg)	Covariate	NA	-0.02 (0.02)	-0.06 (0.08)	-0.06 (0.07)
	Time*	NA	**0.29 (0.08)**	**-0.87 (0.42)**	**0.61 (0.31)**
	Interaction	NA	0.01 (0.01)	-0.03 (0.04)	-0.02 (0.04)
Blood Urea Nitrogen (mg/dl)	Covariate	**0.09 (0.02)**	**0.23 (0.03)**	**-0.21 (0.09)**	NA
	Time*	**0.23 (0.12)**	**0.29 (0.08)**	**-0.87 (0.42)**	NA
	Interaction	**-0.02 (0.01)**	**-0.03 (0.01)**	**0.09 (0.04)**	NA
Serum Glucose (mg/dl)	Covariate	NA	NA	NA	0.01 (0.010
	Time*	NA	NA	NA	**1.86 (0.44)**
	Interaction	NA	NA	NA	**-0.01 (0.00)**
Bicarbonate (mEq/L)	Covariate	-0.08 (0.6)	-0.04 (0.9)	-0.14 (0.24)	NA
	Time*	0.22 (0.12)	**0.29 (0.08)**	**-0.88 (0.42)**	NA
	Interaction	-0.01 (0.03)	0.04 (0.03)	-0.10 (0.11)	NA
Albumin (g/dl)	Covariate	0.63 (0.6)	0.86 (0.77)	**5.55 (1.70)**	-2.25 (1.83)
	Time*	0.23 (0.12)	**0.29 (0.08)**	**-0.89 (0.41)**	**0.65 (0.31)**
	Interaction	-0.62 (0.35)	-0.19 (0.22)	**-2.00 (0.84)**	-0.51 (0.98)
Phosphate	Covariate	0.59 (0.30)	0.28 (0.46)	**-4.23 (1.28)**	NA

(mg/dl)	Time*	**0.24 (0.12)**	**0.29 (0.08)**	**-0.88 (0.42)**	NA
	Interaction	-0.24 (0.18)	0.10 (0.14)	**1.52 (0.61)**	NA
Total Cholesterol (mg/dl)	Covariate	-0.01 (0.00)	0.01 (0.01)	-0.02 (0.01)	0.03 (0.02)
	Time*	**0.25 (0.12)**	**0.29 (0.08)**	**-0.92 (0.42)**	**0.63 (0.31)**
	Interaction	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	-0.01 (0.01)
Hemoglobin (g/dl)	Covariate	-0.02 (0.12)	NA	**2.07 (0.44)**	NA
	Time*	0.22 (0.12)	NA	**-0.98 (0.42)**	NA
	Interaction	0.02 (0.07)	NA	-0.35 (0.21)	NA
White Blood Cell Count (K/mm^3)	Covariate	0.10 (0.11)	NA	0.00 (0.00)	NA
	Time*	**0.24 (0.12)**	NA	**-0.89 (0.42)**	NA
	Interaction	-0.07 (0.07)	NA	0.00 (0.00)	NA
Urine Volume (ml/d)	Covariate	**0.00 (0.00)**	0.72 (0.42)	0.00 (0.00)	0.00 (0.00)
	Time*	**0.25 (0.12)**	0.10 (0.10)	**-0.86 (0.42)**	0.05 (0.36)
	Interaction	0.00 (0.00)	0.01 (0.11)	0.00 (0.00)	0.00 (0.00)
Urine Creatinine (mg/d)	Covariate	**0.00 (0.00)**	**2.56 (0.64)**	**0.00 (0.00)**	0.00 (0.00)
	Time*	**0.25 (0.12)**	0.09 (0.10)	**-0.95 (0.42)**	0.21 (0.36)
	Interaction	0.00 (0.00)	0.02 (0.17)	0.00 (0.00)	0.00 (0.00)
Urine Protein (g/d)	Covariate	0.10 (0.13)	**1.20 (0.17)**	-1.73 (1.01)	-0.15 (0.72)
	Time*	**0.48 (0.15)**	0.13 (0.13)	**-9.29 (3.57)**	-0.89 (1.04)
	Interaction	**0.18 (0.07)**	-0.09 (0.05)	**1.12 (0.47)**	0.65 (0.42)
Urine Phosphate (mg/d)	Covariate	0.00 (0.00)	NA	0.00 (0.00)	NA
	Time*	**0.25 (0.12)**	NA	-0.28 (0.75)	NA
	Interaction	0.00 (0.00)	NA	0.00 (0.00)	NA
Urine Urea Nitrogen (g/d)	Covariate	**0.31 (0.08)**	**0.31 (0.11)**	-0.16 (0.13)	NA
	Time*	**0.25 (0.12)**	0.09 (0.10)	-0.21 (0.58)	NA
	Interaction	-0.04 (0.04)	0.00 (0.00)	0.08 (0.05)	NA
Age (>65 years)	Covariate	0.78 (0.52)	-0.02 (0.53)	NA	NA
	Time*	**0.29 (0.13)**	**0.35 (0.09)**	NA	NA
	Interaction	-0.32 (0.29)	-0.18 (0.17)	NA	NA
Females	Covariate	**-1.43 (0.43)**	**-2.38 (0.54)**	-1.50 (1.81)	1.46 (1.24)
	Time*	0.22 (0.15)	0.09 (0.09)	-0.80 (0.56)	0.11 (0.41)
	Interaction	0.03 (0.24)	**0.52 (0.15)**	-0.18 (0.84)	1.20 (0.61)

* Unit for bias is ml/min/1.73 m^2; Unit for time is ml/min/1.73 m^2 per year

-Variable with p-value for coefficient <0.05 is in bold
Variables that significantly interacted with time in at least in one study are in bold

Although some of these variables are significant, the effects are small. For example, in the MDRD Study a 10 mg/dl higher baseline blood urea nitrogen is associated with a 0.5 ml/min per1.73 m^2 greater change in error at 2 years. In CSG, a 0.5 mg/dl higher baseline albumin is associated with a 0.7 ml/min per 1.73 m^2 greater change in error at 2 years.

6.4 Non-linear Effect of Time – Changing Rate of Change

We examined potential nonlinear trends within and between individuals both graphically by plotting each individual's mean error (difference between measured and estimated GFR) over time as well as an average curve for each study and algebraically by fitting quadratic trends with time.

For one study, this was of the form:

$$Y_{it} = \pi_{i0} + \pi_{i1*}\text{Time}_{it} + \pi_{i2*}Time_{it}^2 + e_{it} \qquad \text{(level 1)} \qquad \text{(A 6.4.1)}$$

$$\pi_{i0} = \gamma_{00} + \zeta_{0i} \qquad \text{(level 2)} \qquad \text{(A 6.4.2)}$$

$$\pi_{i1} = \gamma_{10} + \zeta_{1i} \qquad \text{(A 6.4.3)}$$

$$\pi_{i2} = \gamma_{20} + \zeta_{2i} \qquad \text{(A 6.4.4)}$$

where

$$e_{it} \sim N(0,\sigma_\varepsilon^2) \quad \text{and} \quad \begin{bmatrix}\zeta_{0i}\\ \zeta_{1i}\\ \zeta_{2i}\end{bmatrix} \sim N\left(\begin{matrix}0\\0,\\0\end{matrix}\begin{bmatrix}\sigma_0^2 & \sigma_{01} & \sigma_{02}\\ \sigma_{10} & \sigma_1^2 & \sigma_{12}\\ \sigma_{20} & \sigma_{21} & \sigma_2^2\end{bmatrix}\right)$$

σ_0^2, σ_1^2 and σ_2^2 summarize the between individual variability in initial status, rates of change and the curvature which was specified as the quadratic term for time, respectively.

The fits of the mixed models indicated that a linear relationship was appropriate for MDRD and CSG but in AASK and DCCT, the large value of the estimate relative to the standard error for the quadratic terms indicated a curvature (Appendix 6.4: Table 1).

Appendix 6.4, Table 1: Non-linear relationship of error with time

Coefficient	**Pooled**	**MDRD Study**	**AASK**	**CSG**	**DCCT**
	Coefficient (SE)	*Coefficient (SE)*	*Coefficient (SE)*	*Coefficient (SE)*	*Coefficient (SE)*
Intercept*	-0.0 (0.3)	-2.3 (0.2)	-3. 8 (0.3)	-1.8 (0.9)	5.1 (0.6)
Time*	-1.1 (0.2)	-0.2 (0.4)	-0.2 (0.2)	-0.1 (1.2)	-3.6 (1.7)
Time^2	0.3 (0.05)	0.2 (0.2)	0.1 (0.04)	-0.2 (0.4)	1.3 (0.5)

* Unit for mean error is ml/min/1.73 m^2; Unit for time is ml/min/1.73 m^2 per year

Appendix 6.4, Figure 1: Non-linear Relationship of Mean Error with Time – by Study

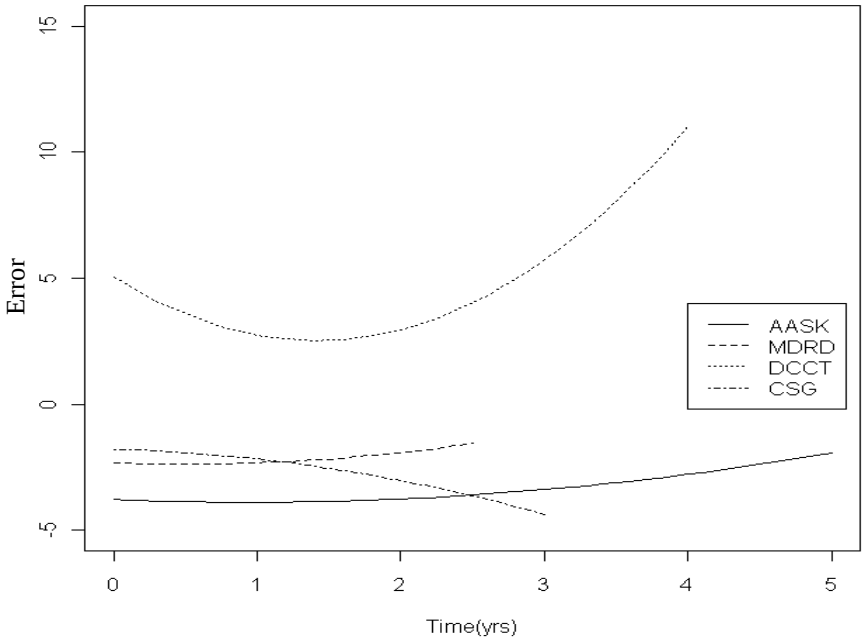


Figure 1 above, shows that the curvature in AASK, while significant, is clinically negligible and can be ignored. We also discounted the non-linear effect in DCCT because of the few measurements available at years 3 and 4 (Table 2).

6.5 Complete Case Analysis

We performed complete case analysis to compare the rate of change in measured and estimated GFR and rate of change in error among those who completed the studies with those who did not. MDRD Study had 196 patients, CSG had 136 patients and DCCT had 151 patients who had follow-up at all the time points in the studies. In AASK, for complete case analysis we included 169 patients who had completed their follow-up until 4 years, as only 86 patients concluded the entire study period of 5 years.

When compared to data from all patients including those who did not finish the study (Table 3), the table below shows that those who completed the study generally had smaller decreases in measured and estimated GFR. Change in error was similar in MDRD Study, AASK and CSG; however in DCCT there was a greater difference in the change in error between those who completed the study versus those who did not. Surprisingly, this difference was caused by the greater decrease in estimated GFR among those who completed the study. This was analyzed further by looking at the distribution of creatinine, age, gender and race among those completed and those who did not (Appendix 6.5, Table 2). Both groups had similar kidney function measures and clinical characteristics. However, non-completers had a slightly lower measured and estimated GFR and serum creatinine. Since only about 10% of the individuals are observed at 4 years, this is not representative of the entire DCCT population.

Appendix 6.5, Table 1: Complete Case Analysis of Measured and Estimated GFR and Difference (error) at First Visit and Rate of Change over Time, by Study

	Population (N), Time (yrs)

Kidney Function Measures		**MDRD Study (196, 2.5 yrs)** *Coefficient (SE)*	**AASK (169, 4 yrs)** *Coefficient (SE)*	**CSG (136, 3 yrs)** *Coefficient (SE)*	**DCCT (151, 4 yrs)** *Coefficient (SE)*
Measured GFR	Mean*	34 (0.9)	49 (1.0)	82 (2.6)	127 (1.6)
	Change**	-2.4 (0.2)	-1.2 (0.2)	-4.8 (0.6)	-0.4 (0.5)
Estimated GFR	Mean*	37 (0.9)	53 (1.2)	82 (2.1)	123 (1.0)
	Change**	-2.6 (0.2)	-1.1 (0.2)	-3.7 (0.6)	-2.2 (0.2)
Error	Mean*	-2.4 (0.4)	-4.2 (0.7)	-0.0 (1.3)	3.9 (1.5)
	Change**	0.2 (0.1)	-0.1 (0.2)	-1.2 (0.5)	1.8 (0.5)

*Unit for error is ml/min/1.73m^2; **Unit for change is ml/min/1.73m^2 per year
Bold indicates p-value <0.05 for the coefficient

Appendix 6.5, Table 2: Exploration of the Kidney Function Measures in DCCT among Completers and Non-completers

Time in years		**Baseline**		**3**		**4**	
		N	***Mean (SD)***	***N***	***Mean (SD)***	***N***	***Mean (SD)***
Completers	Scr	151	0.81 (0.14)	151	0.87 (0.15)	151	0.87 (0.15)
	mGFR	151	127.01 (18.87)	151	123.65 (22.44)	151	126.26 (23.23)
	eGFR	151	122.89 (12.76)	151	115.61 (12.59)	151	114.54 (13.22)
	Age	151	27.11 (7.15)	151	30.11 (7.15)	151	31.11 (7.15)
	Female	151	0.43 (0.50)	151	0.43 (0.50)	151	0.43 (0.50)
	Black	151	0.05 (0.21)	151	0.05 (0.21)	151	0.05 (0.21)
Non-Completers	Scr	1377	0.75 (0.13)	501	0.77 (0.13)	151	0.78 (0.14)
	mGFR	1378	125.58 (22.52)	501	122.62 (22.02)	151	126.26 (23.23)
	eGFR	1377	120.54 (13.65)	501	116.31 (12.31)	151	114.54 (13.22)
	Age	1381	26.82 (7.08)	501	30.61 (6.68)	151	31.11 (7.15)
	Female	1381	0.46 (0.50)	501	0.46 (0.50)	151	0.46 (0.50)
	Black	1381	0.03 (0.18)	501	0.03 (0.18)	151	0.03 (0.18)

Scr-Serum creatinine, mGFR-measured GFR, eGFR-estimated GFR. Age, Female (gender), Black (race) are the other variables in the GFR estimating equation

References

1. Ma, Y.C., et al., *Modified glomerular filtration rate estimating equation for Chinese patients with chronic kidney disease.* J Am Soc Nephrol, 2006. **17**(10): p. 2937-44.

2. Imai, E., et al., *Estimation of glomerular filtration rate by the MDRD study equation modified for Japanese patients with chronic kidney disease.* Clin Exp Nephrol, 2007. **11**(1): p. 41-50.

3. Perrone, R.D., N.E. Madias, and A.S. Levey, *Serum creatinine as an index of renal function: new insights into old concepts.* Clin Chem, 1992. **38**(10): p. 1933-53.

4. Stevens, L.A., et al., *Assessing kidney function--measured and estimated glomerular filtration rate.* N Engl J Med, 2006. **354**(23): p. 2473-83.

5. Levey, A.S., et al., *Using standardized serum creatinine values in the modification of diet in renal disease study equation for estimating glomerular filtration rate.* Ann Intern Med, 2006. **145**(4): p. 247-54.

6. Myers, G.L., et al., *Recommendations for improving serum creatinine measurement: a report from the Laboratory Working Group of the National Kidney Disease Education Program.* Clin Chem, 2006. **52**(1): p. 5-18.

7. Miller, W.G., *Reporting estimated GFR: a laboratory perspective.* Am J Kidney Dis, 2008. **52**(4): p. 645-8.

8. Levey, A.S., et al., *A new equation to estimate glomerular filtration rate.* Ann Intern Med, 2009. **150**(9): p. 604-12.

9. Malamos, B., et al., *125I-sodium iothalamate in the determination of the glomerular filtration rate.* Nucl Med (Stuttg), 1967. **6**(3): p. 304-10.

10. Stevens, L.A., et al., *Impact of creatinine calibration on performance of GFR estimating equations in a pooled individual patient database.* Am J Kidney Dis, 2007. **50**(1): p. 21-35.

11. Stevens, L.A., et al., *Factors other than glomerular filtration rate affect serum cystatin C levels.* Kidney Int, 2009. **75**(6): p. 652-60.

12. Stevens, L.A., Y. Zhang, and C.H. Schmid, *Evaluating the performance of equations for estimating glomerular filtration rate.* J Nephrol, 2008. **21**(6): p. 797-807.

13. *K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification.* Am J Kidney Dis, 2002. **39**(2 Suppl 1): p. S1-266.

14. Waikar, S.S., V.S. Sabbisetti, and J.V. Bonventre, *Normalization of urinary biomarkers to creatinine during changes in glomerular filtration rate.* Kidney Int, 2010. **78**(5): p. 486-94.

15. Scott, P.J. and P.J. Hurley, *Demonstration of individual variation in constancy of 24-hour urinary creatinine excretion.* Clin Chim Acta, 1968. **21**(3): p. 411-4.

16. Greenblatt, D.J., et al., *Variability of 24-hour urinary creatinine excretion by normal subjects.* J Clin Pharmacol, 1976. **16**(7): p. 321-8.

17. Newman, D.J., et al., *Urinary protein and albumin excretion corrected by creatinine and specific gravity.* Clin Chim Acta, 2000. **294**(1-2): p. 139-55.

18. Levey, A.S., et al., *Glomerular filtration rate measurements in clinical trials. Modification of Diet in Renal Disease Study Group and the Diabetes Control and Complications Trial Research Group.* J Am Soc Nephrol, 1993. **4**(5): p. 1159-71.

19. Smith, H.W., *Comparative physiology of the kidney* The Kidney: Structure and Function in Health and Disease: New York, Oxford Univ Press.

20. Perrone, R.D., et al., *Utility of radioisotopic filtration markers in chronic renal insufficiency: simultaneous comparison of 125I-iothalamate, 169Yb-DTPA, 99mTc-DTPA, and inulin. The Modification of Diet in Renal Disease Study.* Am J Kidney Dis, 1990. **16**(3): p. 224-35.

21. Kwong, Y.T., et al., *Imprecision of urinary iothalamate clearance as a gold-standard measure of GFR decreases the diagnostic accuracy of kidney function estimating equations.* Am J Kidney Dis, 2010. **56**(1): p. 39-49.

22. Agarwal, R., *Ambulatory GFR measurement with cold iothalamate in adults with chronic kidney disease.* Am J Kidney Dis, 2003. **41**(4): p. 752-9.

23. Xie, D., et al., *A comparison of change in measured and estimated glomerular filtration rate in patients with nondiabetic kidney disease.* Clin J Am Soc Nephrol, 2008. **3**(5): p. 1332-8.

24. Wang, X., et al., *Validation of creatinine-based estimates of GFR when evaluating risk factors in longitudinal studies of kidney disease.* J Am Soc Nephrol, 2006. **17**(10): p. 2900-9.

25. Stevens, L.A. and A.S. Levey, *Measured GFR as a confirmatory test for estimated GFR.* J Am Soc Nephrol, 2009. **20**(11): p. 2305-13.

26. Cockcroft, D.W. and M.H. Gault, *Prediction of creatinine clearance from serum creatinine.* Nephron, 1976. **16**(1): p. 31-41.

27. Klahr, S., et al., *The effects of dietary protein restriction and 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): p. 877-84.

28. Levey, A.S., et al., *Effects of dietary protein restriction on the progression of advanced renal disease in the Modification of Diet in Renal Disease Study.* Am J Kidney Dis, 1996. **27**(5): p. 652-63.

29. Gassman, J.J., et al., *Design and statistical aspects of the African American Study of Kidney Disease and Hypertension (AASK).* J Am Soc Nephrol, 2003. **14**(7 Suppl 2): p. S154-65.

30. Agodoa, L.Y., et al., *Effect of ramipril vs amlodipine on renal outcomes in hypertensive nephrosclerosis: a randomized controlled trial.* JAMA, 2001. **285**(21): p. 2719-28.

31. Lewis, E.J., et al., *The effect of angiotensin-converting-enzyme inhibition on diabetic nephropathy. The Collaborative Study Group.* N Engl J Med, 1993. **329**(20): p. 1456-62.

32. *The Diabetes Control and Complications Trial (DCCT). Design and methodologic considerations for the feasibility phase. The DCCT Research Group.* Diabetes, 1986. **35**(5): p. 530-45.

33. The Diabetes Control and Complications Trial Research Group (DCCT), *The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus.* N Engl J Med, 1993. **329**(14): p. 977-86.

34. *Diabetes Control and Complications Trial (DCCT): results of feasibility study. The DCCT Research Group.* Diabetes Care, 1987. **10**(1): p. 1-19.

35. Singer, J.D., Willett, J. B., *Applied Longitudinal Data Analysis:Modeling Change and Event Occurrence*. 1st ed. 2003: New York: Oxford University Press.