Item S1. 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 [29]. In the mixed model, the random intercepts and slopes were assumed to be normally distributed about a common mean intercept and slope with a common variance-covariance matrix. The variance of the model's residual at time *t* has a quadratic dependence on time and the covariance has a band diagonal structure such that the covariance between any two pairs of residuals will decline as they become farther apart in time.

The matrix of variance-covariance of the random effects from the mixed model with random intercept and time for the error on the natural scale is given by $\begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{10} & \sigma_1^2 \end{bmatrix} = \begin{bmatrix} 197.0 & -26.6 \\ -26.6 & 10.7 \end{bmatrix}$, where σ_0^2 is the between-subject residual variance in the error at time 0, σ_1^2 is the between-subject residual variance in the error rate of change and $\sigma_{01} = \sigma_{10}$ is the residual covariance between the error at time 0 and the rate of change. The level-1 residual variance across all time points for one subject is $\sigma_e^2 = 41.4$. The variance of the error and the covariance between two errors at two distinct follow-up times on the same subject can be derived using the variance components described above.[Give Singer book reference here]
Because of the small number of measurements on many individuals, we focused on linear and quadratic relationships.

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[29]. The means model assumed no relationship with time; the growth model assumed that an individual's mean error was 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 ith individual at time t takes the form:

$$Y_{it} = \pi_{i0} + e_{it} \quad \text{(level 1)} \qquad \text{(A1.1)}$$

$$\pi_{i0} = \gamma_{00} + \zeta_{0i} \quad \text{(level 2)} \qquad \text{(A1.2)}$$

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} deviated from the ith individual's true error π_{i0} by the within-individual residual e_{it} which had mean 0 and variance σ_e^2 that described the scatter of the individual time-specific errors around their own mean. For person i, the true individual specific mean (π_{i0}) deviated from the population average true mean γ_{00} by the level-2 residual ζ_{0i} which had mean 0 and variance σ_0^2, the between scatter of individual-specific means around the population mean).

In the growth model, time was inserted as a predictor in level-1. We first included no substantive predictors at level-2, so comparison of the growth and means models evaluated how time can explain within-individual variation. The growth model had the form:

$$Y_{it} = \pi_{i0} + \pi_{i1} time_{it} + e_{it} \quad \text{(level 1)} \qquad \text{(A1.3)}$$

$$\pi_{i0} = \gamma_{00} + \zeta_{0i} \quad \text{(level 2)} \qquad \text{(A1.4)}$$

$$\pi_{i1} = \gamma_{10} + \zeta_{1i}. \qquad \text{(A1.5)}$$

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

$$\begin{bmatrix} \zeta_{0i} \\ \zeta_{1i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{10} & \sigma_1^2 \end{bmatrix} \right) \qquad \text{(A1.6)}$$

so that there were variance components for between-individual intercepts and slopes as well as a correlation between the two. Now, σ_e^2 was the within-individual residual variance that summarized the scatter of the errors around the linear change trajectories.

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 ith individual at time t as a function of time as in equation A1.3 above.

Each individual's intercepts and slopes were modeled as a function of study (with DCCT serving as the reference study) as:

$$\pi_{i0} = \gamma_{00} + \gamma_{01} MDRD_i + \gamma_{02} AASK_i + \gamma_{03} CSG_i + \zeta_{0i} \qquad \text{(A2.1)}$$

$$\pi_{i1} = \gamma_{10} + \gamma_{11} MDRD_i + \gamma_{12} AASK_i + \gamma_{13} CSG_i + \zeta_{1i} \qquad \text{(A2.2)}$$

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 had 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, we obtained:

Supplemental Table A: Changes by Study

	Baseline Mean Error (γ_{00})* Coefficient (SE)	Slope/ Rate of Change in Error (γ_{10})** Coefficient (SE)
Pooled Model	0.14 (0.25)	-0.09 (0.08)
Combined Model+		
MDRD Study γ_{01},γ_{11}	**-2.08 (0.47)**	-0.02 (0.14)
AASK γ_{02}, γ_{12}	**-4.09 (0.42)**	**0.30 (0.12)**
CSG γ_{03}, γ_{13}	**-2.01 (0.76)**	**-0.87 (0.28)**
DCCT γ_{00}, γ_{10}	**6.60 (0.42)**	0.09 (0.16)

* 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 486.5with 6 degrees of freedom, (p <0.0001)

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:

$$Y_{it} = \pi_{i0} + \pi_{i1} * Time_{it} + \pi_{i2} * Time_{it}^2 + e_{it} \quad \text{(level 1)} \qquad \text{(A 4.1)}$$

$$\pi_{i0} = \gamma_{00} + \zeta_{0i} \quad \text{(level 2)} \qquad \text{(A 4.2)}$$

$$\pi_{i1} = \gamma_{10} + \zeta_{1i} \qquad \text{(A 4.3)}$$

$$\pi_{i2} = \gamma_{20} + \zeta_{2i} \qquad \text{(A 4.4)}$$

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 while the average trends within studies were linear, significant nonlinearity was apparent for individual subjects in their rate of change (Supplemental Table B), Because these individual nonlinearities showed no consistent pattern, we chose to stick with the linear model overall and for each study.

Supplemental Table B: 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*	0.35 (0.26)	**-2.23 (0.20)**	**-3.87 (0.27)**	-1.82 (0.96)	**6.67 (0.62)**
Time*	**-0.94 (0.18)**	0.29 (0.18)	0.01 (0.18)	-1.53 (1.10)	-0.79 (0.81)
Time^2	**0.23 (0.04)**	-0.08 (0.05)	0.06 (0.04)	0.28 (0.32)	0.23 (0.20)

* Unit for mean error is ml/min/1.73 m^2; Unit for time is ml/min/1.73 m^2 per year
Bolded terms are significant. By the likelihood ratio test, the difference and p-value between the linear time and the quadratic time models with 4 degrees of freedom in the pooled, MDRD, AASK, CSG and DCCT were (525.6, p <0.0001), (77.1, p <0.0001), (54.7, p <0.0001) and (7.8, p =0.10) respectively

.