Appendix I: Full Methods

Participants and Design

The Chronic Kidney Disease in Children (CKiD) study enrolled children aged 1 to 16 years with a Schwartz-estimated GFR [14-16] between 30 and 90 ml/min|1.73 m^2 who had never been dialyzed or undergone organ transplant. Participants were drawn from 43 participating tertiary care pediatric nephrology programs across the U.S. and 2 sites in Canada. The baseline study visits occurred between January 19, 2005 and August 3, 2009, with annual follow -up visits. Institutional Review Boards for each participating site approved the study protocol and the study has been described previously [13].

At the baseline and first annual one year follow-up visits, GFR was determined by directly measured plasma iohexol (GE Healthcare, Amersham Division, Princeton, NJ) disappearance curves; details of the GFR assessment methods have been published previously.[18] An estimated GFR value was used when a directly measured value was unavailable [17]. Basic metabolic profile, including measurement of creatinine, was assessed using an enzymatic method on the Bayer Advia 2400 analyzer (Siemens Diagnostics, Tarrytown NY).

For the present study we nested a case-control design within the CKiD cohort by matching children who had been observed to initiate dialysis or undergo kidney transplant (renal replacement therapy [RRT] cases) to children who, at the time of the case event, had not yet experienced an RRT event. Controls, however, could become cases in the study at a later time point. Thus the design matched cases to controls on time on study. We also matched on CKD stage at baseline and glomerular/non-glomerular diagnosis. Cases were matched to one control without replacement such that each case had a unique control and cases for which an appropriate match was unavailable were removed from the analysis. The order of matching was determined by a random computer generated sequence.

Statistical Analysis

GFR trajectories of cases and controls were modeled using a log-linear mixed effects model of the form

$$\log(GFR_{ij}) = (\alpha_0 + a_i) + \alpha_1 case_i + \alpha_2 age_i + \alpha_3 race_i + \alpha_4 sex_i + \alpha_5 proteinuria\ status_i + (\beta_0 + b_i)t_{ij} + \beta_1 t_{ij} case_i + \varepsilon_{ij},$$

where β_0 is the rate of log(GFR) change in controls, and β_1 is difference in the rates of change of log(GFR) comparing cases and controls. The intercept, α_0, represented the log(GFR) at the time of RRT for the case (or comparable time from baseline for the controls) and time proceeded negatively in years such that -1 represented 1 year prior to RRT. Thus all case GFR trajectories were anchored at RRT with their matched control GFR trajectories anchored at the same time from baseline as the respective cases. The model included subject-specific random effects for the intercept and slope (a_i and b_i) distributed according to $\begin{pmatrix} a_i \\ b_i \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_1^2 & \rho\tau_1\tau_2 \\ & \tau_2^2 \end{pmatrix}\right]$. Finally ε_{ij} represented the random error with a distribution of $\varepsilon_{ij} \sim N\ (0, \sigma^2)$. The model was adjusted for age (continuous), race (white versus nonwhite), sex (male versus female) and baseline proteinuria status (categorized as urine protein to creatinine ratio<0.2, 0.2 to 2.0, and ≥2.0).

To address the study hypothesis that cases experienced an acceleration of their kidney function decline prior to RRT, the above model was refined to allow a change in slope (spline)

proximate to the RRT event. The spline term in this piecewise log-linear mixed effects was positioned at 18 months prior to the date of RRT (t = -1.5). The model was of the form

$$\log(GFR_{ij}) = (\alpha_0 + a_i) + \alpha_1 case_i + +\alpha_2 age_i + \alpha_3 race_i + \alpha_4 sex_i + \alpha_5 proteinuria\ status$$
$$+(\beta_0 + b_i)t_{ij} + \beta_1 t_{ij} case_i + \beta_2\left(t_{ij} - (-1.5)\right)_{t-} + \beta_3\left(t_{ij} - (-1.5)\right)_{t-} case_i + \varepsilon_{ij}$$

with a slope parameter for the spline term (β_2) and the interaction between case status and the spline (β_3) added to the log-linear mixed effects model to form the piecewise log-linear model. The breakpoint for the change in slope was set at 18 months prior to RRT to balance the desire to look for slope changes proximate to RRT with the limitations of the data and the need to have at least two GFR measurements from most participants to robustly fit a slope representative of the individual GFR trajectories.

Akaike Information Criterion (AIC) was used to assess the fit of both models described above. All of the analyses were conducted using STATA/MP 11.2 (Statacorp LP) and the graphics were produced using SAS 9.3 (SAS Institute, Inc).

Appendix II. Coefficient estimates from the log-linear and piecewise log-linear model

Log-linear model:

$$lGFR_{ij} = (\alpha_0 + a_i) + \alpha_1 case_i + \alpha_2 age_i + \alpha_3 race_i + \alpha_4 sex_i + \alpha_5 proteinuria\ status_i + (\beta_0 + b_i)t_{ij} + \beta_1 t_{ij} case_i + \varepsilon_{ij},$$

Piecewise log-linear model:

$$lGFR_{ij} = (\alpha_0 + a_i) + \alpha_1 case_i + + \alpha_2 age_i + \alpha_3 race_i + \alpha_4 sex_i + \alpha_5 proteinuria\ status + (\beta_0 + b_i)t_{ij} + \beta_1 t_{ij} case_i + \beta_2 \left(t_{ij} - (-1.5)\right)_{t-} + \beta_3 \left(t_{ij} - (-1.5)\right)_{t-} case_i + \varepsilon_{ij}$$

	Log-linear Mixed Effects Model*			Piecewise log-linear Mixed Effects Model¶		
Parameter	**Coefficient Estimate**	**SE**	**P-value**	**Coefficient Estimate**	**SE**	**P-value**
α_0	3.493	0.072	<0.001	3.456	0.074	<0.001
α_1	-0.515	0.033	<0.001	-0.570	0.040	<0.001
α_2	0.010	0.005	0.033	0.009	0.005	0.068
α_3	0.070	0.038	0.061	0.069	0.037	0.061
α_4	-0.067	0.036	0.065	-0.064	0.036	0.074
α_5	-0.034	0.028	0.228	-0.038	0.028	0.168
β_0	-0.032	0.012	0.009	-0.094	0.028	0.001
β_1	-0.169	0.015	<0.001	-0.297	0.033	<0.001
β_2	-	-	-	0.097	0.037	0.008
β_3	-	-	-	0.224	0.046	<0.001
AIC		260.78			149.14	

*In the log-linear mixed effects model:

Slope for controls: β_0=-0.032 (SE=0.012; P-value=0.009)

Slope for cases: $\beta_0 + \beta_1$=-0.201 (SE=0.011; P-value<0.001)

Difference in slopes comparing cases and controls: β_1 = -0.169 (SE=0.015; p-value<0.001)

¶In the piecewise log-linear mixed effects model:

Slope before -1.5 years for controls: $\beta_0 + \beta_2$=0.003 (SE=0.015; P-value=0.866)

Slope after -1.5 years for controls: β_0=-0.094 (SE=0.028; P-value<0.001)

Slope before -1.5 years for cases: $\beta_0 + \beta_1 + \beta_2 + \beta_3$=-0.070 (SE=0.015; P-value<0.001)

Slope after -1.5 years for cases: $\beta_0 + \beta_1$=-0.391 (SE=0.019; P-value<0.001)

Difference in slopes comparing cases and controls before -1.5 years: $\beta_1 + \beta_3$ =-0.073 (SE=0.021; P-value<0.001)

Difference in slopes comparing cases and controls after -1.5 years: β_1 =-0.297 (SE=0.033; P-value<0.001)

Difference in early and late slopes of controls: - β_2=-0.097 (SE=0.037; P-value=0.008)

Difference in early and late slopes of cases: - β_2- β_3=-0.320 (SE=0.028; P-value<0.001)