Modeling GFR trajectories in diabetic nephropathy

Kevin V. Lemley,1,2,a Derek B. Boothroyd,3,a Kristina L. Blouch,1 Robert G. Nelson,4 Lois I. Jones,1 Richard A. Olshen,3 and Bryan D. Myers1

Departments of 1Medicine, 2Pediatrics, and 3Health Research and Policy, Stanford University School of Medicine, Stanford, California; and 4Phoenix Epidemiology and Clinical Research Branch, National Institute of Diabetes and Digestive and Kidney Diseases, Phoenix, Arizona

Submitted 1 March 2004; accepted in final form 9 May 2005

Lemley, Kevin V., Derek B. Boothroyd, Kristina L. Blouch, Robert G. Nelson, Lois I. Jones, Richard A. Olshen, and Bryan D. Myers. Modeling GFR trajectories in diabetic nephropathy. Am J Physiol Renal Physiol 289: F863–F870, 2005. First published May 17, 2005; doi:10.1152/ajprenal.00068.2004.—In an 8-year longitudinal study of Pima Indians with type 2 diabetes and nephropathy, we used statistical techniques that are novel and depend on minimal assumptions to compare longitudinal measurements of glomerular filtration rate (GFR). Individuals enrolled with new-onset microalbuminuria either progressed to macroalbuminuria (progressors, n = 13) or did not progress (nonprogressors, n = 13) during follow-up. Subjects with new-onset macroalbuminuria at screening were also followed (n = 22). Patients had their GFR determined serially by urinary iothalamate clearances (average 11 clearances; range 6–19). GFR courses of individuals were modeled using an adaptation of smoothing and regression cubic B-splines. Group comparisons were based on five-component vectors of fitted GFR values using a permutation approach to a Hotelling’s T2 statistic. GFR profiles of initially microalbuminuric progressors differed significantly from those of nonprogressors (P = 0.003). There were no significant baseline differences between progressors and nonprogressors with respect to any measured clinical parameters. The course of GFR in the first 4 yr following progression to macroalbuminuria in initially microalbuminuric subjects did not differ from that in newly screened macroalbuminuric subjects (P = 0.27). Without imposing simplifying models on the data, the statistical techniques used demonstrate that the courses of decline of GFR in definable subgroups of initially microalbuminuric diabetic Pima Indians, although generally progressive, follow distinct trajectories that are related to the extent of glomerular barrier dysfunction, as reflected by the evolution from microalbuminuria to macroalbuminuria.

Pima Indian; smoothing splines; iothalamate clearance; permutation testing; progression

The natural history of diabetic nephropathy is generally characterized by a variable period of hyperfiltration followed by an inexorably progressive decline of glomerular filtration rate (GFR) once overt proteinuria appears (17). Variations in the rate of GFR decline in patients with type 2 diabetic nephropathy have been described before in studies with short follow-up periods (19). The Pima Indians of Arizona have a very high incidence rate of type 2 diabetes mellitus and an inordinately high risk of end-stage kidney failure due to diabetic nephropathy. Considerable heterogeneity in the actual evolution of diabetic nephropathy, however, is present in this population. For example, we reported earlier that the pattern of progression from microalbuminuria to macroalbuminuria may be quite variable (14).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
end-stage renal failure). Macroalbuminuria and microalbuminuria were defined based on results of the geometric mean of the albumin/creatinine ratio (ACR) in three screening urine samples together with the ACR in two subsequent samples following enrollment. Individuals were classified on the basis of having at least two of these three values for urinary ACR within the microalbuminuric (30–299 mg/g) or macroalbuminuric (≥300 mg/g) range. Particular attention was paid to examining all available records of quantitative proteinuria or albuminuria of the enrolled subjects, so as to exclude any individuals in whom prior evidence of either microalbuminuria or macroalbuminuria was found. The protocols were approved by the institutional review boards of Stanford University and the Gila River Indian Community Tribal Council.

During this extended follow-up, 13 of the initially microalbuminuric subjects progressed to macroalbuminuria and 13 remained persistently microalbuminuric. The approximate date of progression was determined from the time course of ACR values by the consensus of three observers. This approach was necessitated by the significant variability often seen in the course of ACR in such subjects (14).

The GFR course after progression to macroalbuminuria were examined in 11 of the initially microalbuminuric subjects who had at least 4 yr of follow-up after progression. In addition to these 11 progressors, we studied the postprogression courses in 7 microalbuminuric subjects, whose ascertainment with microalbuminuria preceded enrollment in the original study, but who progressed to persistent macroalbuminuria under observation. Only the macroalbuminuric stages of these individuals’ GFR courses were compared in the present study. Thus there were two partially overlapping groups of microalbuminuric subjects: 26 in whom the entire GFR course from first detection of microalbuminuria was studied and a group of 18 subjects (11 from the former 26 plus 7 others) in whom only the GFR course after progression to macroalbuminuria was analyzed.

Study protocol. The original 4-yr longitudinal study of Pima Indian subjects with microalbuminuria or macroalbuminuria (17) envisaged physiological studies [GFR, renal plasma flow (RPF), etc.] occurring every 6 mo. This frequency of sampling was accomplished in about one-half of the patients for the first 4 yr. In the remaining one-half of the subjects, the first follow-up study was 1 yr after the baseline studies, with the study frequency increasing to every 6 mo thereafter. Although many of the participants continued to have regular physiological studies as a part of a long-term extension of the original 4-yr protocol, many of the microalbuminuric subjects had a hiatus of ~2 yr in their physiological studies before recommencing intensified follow-up for a final 3–4 yr. Hence the sampling of physiological data was spaced irregularly. The resulting difficulties with tracking raw data were overcome by suitably adapting the technique of smoothing splines (vide infra). During the entire study, ACR was determined every 3 to 6 mo.

Laboratory methods. GFR was estimated by urinary iothalamate clearance and RPF by PAH clearance. The filtration fraction was calculated as the ratio of GFR to RPF. The plasma osmotic pressure (\(\pi_A\)) was determined by membrane osmometry, and HbA1c by HPLC, as previously described (14). In some cases, values for HbA1c were inferred from electrophoretically measured HbA1, by means of a regression equation based on samples from 133 patients for which both methods were used (18). Serum cholesterol was measured by a colorimetric method. Urinary albumin concentration was measured by nephelometry. Urinary creatinine concentration was determined by a modified Jaffé reaction, as previously described (14).

The iothalamate and PAH clearance studies consisted of four timed urine collections bracketed by five plasma samples starting 1 h after bolus intravenous infusions of iothalamate (300 mg for subjects up to 100 kg body wt, 3 mg/kg for subjects >100 kg body wt) and PAH (8 mg/kg for subjects with serum creatinine concentrations <2.0 mg/dl; 4 mg/kg for subjects with serum creatinine concentrations ≥2.0 mg/dl). The bolus was followed by continuous infusions to maintain plasma levels of 1–2 and 1–3 mg/dl for the two clearance markers, respectively. The clearances were conducted during a sustained diuresis brought about by drinking 10 ml/kg body wt of water (maximum 900 ml) initially and replacing urinary losses throughout the study. A PAH extraction ratio of 0.85 (for the calculation of effective RPF) was used for subjects with a GFR ≥80 ml/min; a ratio of 0.70 was used if GFR was <80 ml/min (2). Baseline values for GFR and RPF were calculated both as absolute values (ml/min) and normalized to 1.73 m² of body surface area. Longitudinal measures of GFR and RPF were not adjusted for body surface area, to avoid changes due only to secular increases in body weight.

Statistics and curve estimation. The GFR courses of the individual patients were estimated from the iothalamate clearance data (with 4 clearance periods per study). An average of 11 complete clearance studies were obtained per patient (range 6–19). The fitting for each individual was from study entry either to the last measured GFR or to the onset of dialysis. Comparisons of the fitted GFR courses of the initially microalbuminuric nonprogressors and progressors were at equally spaced points over 8 yr. GFR courses for progressors, from the point of progression, were compared with those of initially macroalbuminuric subjects, from the time of ascertainment, by evaluating fitted curves at equally spaced points over 4 yr. These are the periods for which complete data were available. Complete details of the statistical analysis are presented in the APPENDIX.

One of the initially microalbuminuric subjects had a GFR course that appeared to deviate significantly from the rest of that cohort. The GFR appeared to increase considerably at the end of follow-up, although there was also a large difference in GFR between the last two measurements. We tested for outlier status of this individual with respect to both the cohort of 13 newly microalbuminuric progressors and the expanded group of 18 microalbuminuric subjects (not all newly detected with microalbuminuria at study entry) who progressed to macroalbuminuria under observation. From the mean and covariance structure of the five fitted GFR values of the cohort, a Hotelling’s T² statistic was developed to compare the apparent outlier with the remainder of the groups (15). The P values for the comparisons were 0.086 for the group of 13 individuals and 0.055 for the group of 18 individuals. We also varied the number of fitted values for which outlier status was evaluated. The P values were very small when a small number of points was used and much larger when more points were used. The borderline statistically significant results of these tests and the appearance of the GFR curve led us to exclude this individual from the final analysis.

For ease of clinical interpretation and due to the variable durations of follow-up, the change in GFR was also expressed as a time-averaged rate of change (ml/min−1·yr−1) in the various groups. These rates were calculated by simple linear regression on the full nine fitted GFR values.

Differences in baseline values of clinical and demographic parameters (age, gender, duration of diabetes, HbA1c, etc.) were evaluated by either t-test or Mann-Whitney U-test for continuous variables and Fisher’s exact test for categorical variables. ACR differences between groups were assessed by t-test on log[ACR]. Data are represented as means ± SD or median (interquartile range).

RESULTS

Glomerular function and structure. The baseline characteristics of the macroalbuminuric and microalbuminuric subjects are reported in Table 1. The groups differed with respect to duration of diabetes, mean arterial pressure (MAP), ACR, body surface area-corrected GFR, filtration fraction, and \(\pi_A\). For comparison, the average GFR in nondiabetic Pima Indian subjects is 123 ± 22 ml/min (17).

The baseline characteristics of the progressors and nonprogressors among the initially microalbuminuric subjects are reported in Table 2. There were no differences in gender mix
Clinical characteristics at the time of progression of the initially microalbuminuric subjects who progressed to macroalbuminuria under observation are reported in Table 3. Compared with the screened macroalbuminuric subjects at baseline, the microalbuminuric progressors at the time of progression had a lower median ACR (403 vs. 904 mg/g, \( P = 0.014 \)) and a higher filtration fraction (0.19 ± 0.02 vs. 0.16 ± 0.04, \( P = 0.025 \)). In contrast, there were no differences with respect to the duration of diabetes, use of ACE inhibitors, age, gender mix, BMI, GFR, RPF, MAP, \( \pi_A \), Hba1c, or serum cholesterol.

**Modeling of GFR course.** Figure 1 shows three examples of raw GFR data (3–4 clearance periods per study) and the fitted GFR curves made using smoothing and regression B-splines as described. A nonlinear GFR trend is apparent in each case illustrated. A substantial intraday variability in GFR among the four collection periods is obvious in some clearance studies, although the overall intraday variability was virtually identical to that of our historical healthy controls and other groups with renal disease (coefficient of variation ~10%).

The fitted individual GFR curves of the 22 macroalbuminuric and 26 microalbuminuric subjects are shown in Fig. 2. It is clear from the figure that almost all of the former had developed kidney failure within 8 yr of entering the study, while only 2 of 26 initially microalbuminuric subjects had gone on to kidney failure over this period. This is despite the fact that the screened macroalbuminuric subjects had only a slightly lower average GFR than the microalbuminuric subjects at the time of enrollment \([136 ± 56 vs. 155 ± 47 ml/min, \( P = \) not significant (NS)]\).

A comparison of the mean fitted GFR curves over 8 yr of follow-up in the initially microalbuminuric subjects shows that those who progressed to persistent macroalbuminuria had a greater rate of decline of GFR than those who remained microalbuminuric (Fig. 3). There was an apparent stabilization in the GFR toward the end of follow-up among the persistently microalbuminuric subjects. The GFR trajectories of these two groups differed significantly as indicated by a comparison of their five-vectors of fitted values using Hotelling’s T2 test \( (P = 0.004) \).
We also compared GFR trajectories over the first 4 yr from the onset of macroalbuminuria in subjects who were found to be macroalbuminuric at screening with those in subjects who progressed to macroalbuminuria after enrollment with microalbuminuria (Fig. 4). Although the GFR curves over 4 yr in these two groups appear to differ, the differences were not statistically significant as indicated by comparison of their five-vectors of fitted values ($P = 0.27$).

Inclusion of the apparent outlier would have increased the significance of the difference between the postprogression course of the initially microalbuminuric subjects and the initially macroalbuminuric subjects. The group difference would still only have approached borderline significance ($P = 0.11$). It would have decreased to borderline significance the difference between the microalbuminuric progressors and nonprogressors.

We found a nonlinear GFR time course in many of our patients (Fig. 2; APPENDIX). To make our results easier to interpret clinically, however, we calculated an average rate of change of GFR for the various groups based on a linear regression on the fitted GFR values. The average rate of decline of GFR over 8 yr of follow-up in the nonprogressors was $4.9 \text{ ml/min/yr}$, whereas in the initially microalbuminuric progressors it was $9.3 \text{ ml/min/yr}$. The average rate of decline of GFR over 4 yr after progression in the course of the initially microalbuminuric subjects and the initially macroalbuminuric subjects.

Fig. 1. Regression and smoothing B-spline fits to the raw glomerular filtration rate (GFR) data of 3 subjects. Left: initially microalbuminuric nonprogressor. Middle: initially microalbuminuric progressor. Right: initially macroalbuminuric (macro) subject. Notice the significant intraday variability in GFR for some of the individual clearance studies.

Fig. 2. GFR courses of 22 macroalbuminuric (red) and 26 initially microalbuminuric (blue) Pima Indians with type 2 diabetes. The GFR courses are nonlinear in many of the subjects. A: entire GFR courses of the subjects. B: entire GFR courses of the initially macroalbuminuric subjects and only the postprogression courses of the initially microalbuminuric subjects.

Fig. 3. Mean GFR courses (fitted values) of initially microalbuminuric subjects who progressed to persistent macroalbuminuria ($n = 13$, ○) or remained microalbuminuric ($n = 13$, ■) over the course of 8 yr of follow-up.
initially microalbuminuric subjects was 10.5 ml·min⁻¹·yr⁻¹, compared with 16.9 ml·min⁻¹·yr⁻¹ in subjects screened with macroalbuminuria.

DISCUSSION

Diabetic nephropathy is the leading cause of kidney failure worldwide. Establishing a comprehensive characterization of the decline of GFR in patients with type 2 diabetic nephropathy is important both for studies of causal factors and of proposed therapies. In addition, defining quantitative differences in the rates of decline of GFR in different populations may help to elucidate potential genetic and environmental risk factors for disease progression. The Pima Indians of Arizona have a very high incidence of type 2 diabetes mellitus and diabetic nephropathy (17). Moreover, compared with European populations, diabetic Pima Indians have an accelerated rate of decline in GFR once they develop overt nephropathy (12). Thus differences in the patterns of loss of GFR among nephropathic Pima Indians may in fact be more pronounced than within these comparison populations.

To study changes in GFR early in the course of diabetic nephropathy requires the use of precise methods such as urinary iothalamate clearances. Because the initial glomerular hyperfiltration characteristic of diabetes results in relatively low serum creatinine concentrations, the earliest declines in GFR will be accompanied by undetectable changes in the serum creatinine concentration (16). Using precise methods, we previously were able to demonstrate relatively small decreases in GFR over just 4 yr in Pima Indians with incipient (microalbuminuric) diabetic nephropathy (13).

Even using precise methods, however, significant variability exists over time in measurements of GFR in diabetic Pima Indians. This variability includes both the usual intraday coefficient of variation of ~10% in iothalamate clearances among the clearance periods, as seen in other populations, and a significant nonlinearity in the time course of GFR in many cases (Figs. 1 and 2). The purpose of the current study was to develop a method of analyzing such longitudinal GFR data by adapting smoothing and regression cubic B-splines, with statistical comparisons being made among the resulting fitted GFR curves based on the permutation distribution of Hotelling’s T² statistic (3). This approach allows us to better incorporate the totality of the individual clearance data as well as to test for differences in GFR trajectories more complex than simple linear trends.

That the smoothing procedures did not significantly distort the underlying data is suggested by the fact that the mean average rate of GFR change based on the raw data did not differ significantly from the mean of the average rate of GFR change based on the fitted data in any of the groups (data not shown). Figures 1 and 2 illustrate that the rate of GFR change is in fact not constant over time, so that simple average rates of GFR change with time will often fail to capture the complexity of the GFR time courses. Although statistical methods based on smoothing and regression B-splines utilize the totality of GFR data to yield more informative GFR-vs.-time curves, their use may introduce challenges with regard to clinical interpretation, inasmuch as the principal outcome variable is a five-vector of fitted GFR values rather than a simple rate of change in GFR. In the current study, we addressed this difficulty by also providing an average rate of GFR change as a single number (ml·min⁻¹·yr⁻¹) based on a linear fit to the data. Statistical tests were, however, performed only on the complete five-vector representations, for the reasons outlined above.

Using these newly adapted statistical methods, we were able to demonstrate that the time courses of GFR in Pima Indians with type 2 diabetes who progress from microalbuminuria to macroalbuminuria during follow-up differ from those of individuals who remain microalbuminuric throughout. Patients with persistent microalbuminuria have a somewhat more benign course, with GFR decreasing by 40 ml/min on average over 8 yr and generally remaining within the normal range at the end of follow-up. It is not clear whether the apparent flattening of the group average GFR curve toward the end of follow-up in the persistently microalbuminuric subjects represents true long-term stabilization (Fig. 3), as fitted values may show some instability at the boundaries of their ranges. There were no significant differences in the baseline clinical parameters describing the initially microalbuminuric patients who followed these two different courses (Table 2), although there were slightly higher values of HbA1c, BMI, MAP, and disease duration in the subjects who progressed to macroalbuminuria during follow-up, suggesting that these subjects had slightly more severe or advanced disease even at the time of ascertainment than those who did not progress. The baseline average GFR and median ACR values in the two groups, on the other hand, were virtually identical (154 ± 27 and 157 ± 63 ml/min for GFR and 61 and 69 mg/g for ACR in progressors and nonprogressors, respectively).

A comparison of clinical characteristics among the progressors and nonprogressors at 48 mo of follow-up showed that for some of those variables in which there was a statistically nonsignificant trend for the future progressors to exceed the nonprogressors at baseline (MAP, HbA1c), statistically significant differences had emerged at 48 mo. Whether these differ-
GFR courses did not differ significantly between subjects who were found at screening to have macroalbuminuria compared with the “postprogression” courses of subjects who had microalbuminuria at screening and who developed macroalbuminuria during follow-up, despite the higher apparent rate of GFR decline in the former (Fig. 4). The average linear rates of decline of GFR differed considerably among the three groups: 4.9 ml·min^{-1}·yr^{-1} in microalbuminuric nonprogressors, 9.3 ml·min^{-1}·yr^{-1} in initially microalbuminuric progressors, and 16.9 ml·min^{-1}·yr^{-1} in the subjects screened with macroalbuminuria. All three groups had average rates of GFR decline considerably greater than those expected in normal individuals over 40 yr old as a result of aging, i.e., −0.5–1 ml·min^{-1}·yr^{-1} (5, 27). In addition, in the latter two groups of Pima Indians, the rate of GFR decline was more than twice that found in studies of European type 2 diabetic patients with macroalbuminuria (4, 19).

As stated above, the GFR time courses were nonlinear in many individuals (Figs. 1 and 2). Three general classes of shapes could be distinguished: an essentially linear decline, a bimodal decline, and a pattern of variable decline. The pattern of linear decline most likely reflects subjects already on the “slippery slope,” in whom compensatory hyperfiltration in the remnant nephrons was already maximal at the time of their entry into the study. In these individuals, all subsequent loss of functioning nephrons was reflected in a loss of GFR. It is interesting that this seemed to occur even in subjects who had GFR values greater than 200 ml/min at study entry. Under the same interpretation, those subjects who had a bimodal decline in GFR probably entered the slippery slope phase during follow-up. Most interesting are those subjects who showed a pattern of deceleration or stabilization of their GFR course during follow-up. A longer period of follow-up will be necessary to determine whether their apparent GFR stability is persistent.

Our interpretation of these GFR time courses is subject to several limitations. Although the initially microalbuminuric subjects were enrolled under consistent entry criteria, those who progressed to macroalbuminuria during follow-up had some clinical characteristics (greater duration of diabetes, BMI, higher HbA1c, greater MAP) suggesting more advanced or severe disease at the time of enrollment. These differences were not statistically significant and seem too small to account for the subsequent development of such significant differences in the course of GFR between the two groups. By 48 mo of follow-up, some of these factors had become significantly different between progressors and nonprogressors, despite the fact that their GFR values were still nearly identical. The role of these covariates in the subsequent GFR course certainly merits further study.

The subjects screened with macroalbuminuria and those who progressed to macroalbuminuria under observation may not have been as comparable as the microalbuminuric progressors and nonprogressors at enrollment. In particular, the precision of the estimation of the time of onset of macroalbuminuria was different in these two groups. The subjects with macroalbuminuria at the original screening may have had their previous ACR up to 2 yr earlier, whereas the subjects screened as microalbuminuric who progressed under observation had an ACR determined every 3–6 mo. We reported before on the considerable heterogeneity in time courses of progression of microalbuminuria to macroalbuminuria in diabetic Pima Indians (14). The combined effects of difficulties in “aligning” the starting points of their macroalbuminuric courses and the nonlinear features of their GFR curves tend to exacerbate the difficulties in comparing their subsequent GFR courses.

We have not formally evaluated the contribution of exposure to ACE inhibitors or angiotensin receptor blockers to the differences in GFR courses in the different groups. Although a slightly greater number of initially microalbuminuric subjects (2/18) were receiving ACE inhibitors at the time of their progression to macroalbuminuria than were receiving them when found to be macroalbuminuric at screening (0/22), four of the latter subjects were taking ACE inhibitors at some point during the first 4 yr of the study. It is unlikely that these low rates of exposure significantly influenced the GFR courses of the groups as a whole. The relatively infrequent use of ACE inhibitors reflects the low incidence of hypertension in the diabetic population of Pima Indians. Only in the last several years of the study has it become a common practice to use ACE inhibitors in normotensive type 2 diabetic individuals.

The lack of a clear association of progression from microalbuminuria to macroalbuminuria during 8 yr of follow-up with traditional clinical variables (gender, age, HbA1c, MAP, etc.) determined at baseline suggests that other factors may influence both the progression to macroalbuminuria and the attendant differences in the courses of GFR decline. Of course, significant differences in MAP and HbA1c between progressors and nonprogressors that had emerged by 48 mo of follow-up may have contributed to the subsequent divergence of their GFR courses. This needs to be addressed in a separate analysis based on a comparison of the time series of these variables.

Our study did not address the role of pertinent genetic factors, such as genotypes of angiotensin-converting enzyme, angiotensinogen, bradykinin 2 receptor, nitric oxide synthase, that may influence characteristics of the progression of diabetic nephropathy. As in some other populations, the insertion/deletion polymorphism of the angiotensin I-converting enzyme gene is associated in Pima Indians with differences in plasma ACE activity (7) and thus might influence the evolution of kidney disease. The ability to precisely quantify changes in GFR over extended periods of time should allow more sensitive investigations of the role of both genetic and environmental factors in disease progression. This may prove particularly valuable in studies conducted within well-characterized and relatively genetically homogenous populations with a high incidence of diabetic nephropathy such as the Pima Indians of Arizona (10). Characterization of GFR courses in terms of five-vectors may appear to complicate an easy physiological interpretation of GFR changes. It should, however, also be amenable to associative study of complex genotype segregation with specific quantitative traits, using modifications of techniques developed for more simple (one dimensional) quantitative traits. Representation of the course of GFR as a five-vector captures more information than the average linear rate of GFR change without imposing an arbitrary structure on the GFR course. We submit that the statistical techniques described above are particularly suitable for analysis of longitudinal studies of irregularly spaced and noisy physiological variables, such as GFR.
APPENDIX

There are problems inherent in our making inferences about individual GFR curves that owe to many sources. One is that subjects entered the study at different stages of the disease. Within each stage, the extent of renal injury varied over a wide range. Another problem is that we had an incomplete set of measurements on each subject. A third is that even after a subject’s entry to the study, measurements for that subject were not made on a regular grid of points. The value of GFR at dialysis or end-stage renal failure (ESRF) is not available by clearance techniques, so we were forced to impute. We used values of 3 and 7 in our analyses, although the exact choices seemed not to matter. To check, we did simulations, taking the macros only because among them were almost all the subjects reaching ESRF. For each such case, we replaced the imputed clearance values with four new imputed values, chosen independently from a uniform distribution on the interval [0, 10]. We then refit the macros and compared the values of the fits at 0, 1, 2, 3, and 4 yr. The majority of the values changed by less than 0.05 (with median change −0.01, median change in absolute value 0.33). More than 80% changed by less than 1.

Our general approach to fitting was that of the functional data analysis of longitudinal curves (21, 22). To begin, the GFR curve for a subject is written as a sum of random subject-specific coefficients times respective B-spline basis functions, plus additive, independent noise. Thus the kth measurement at the jth time for the ith individual on y = GFR is assumed to follow this regression spline model:

$$y_{ijk} = y_j(t_i) = \sum_{m=1}^{M} a_{ij} B_m(t_i) + \epsilon_{ijk}$$  \hfill (A1)

Each $B_m(\cdot)$ is a cubic B-spline. In (A1), for each subject i, the vector $[a_{ij}]$ was fitted. They are assumed random, and we had but a single realization for each subject. We used 21 cubic B-splines, with knots separated by 18 mo on a fixed grid. This entailed that for a given subject i, most cubic spline basis functions were identically 0 over the domain for which GFR was fitted for that subject. The $[\epsilon_{ijk}]$ were assumed independent with common mean 0 and variance $\sigma_j^2$. Independence across k for fixed (i, j) was investigated and found not to be sustainable; but its violation was not important to our inference, either. We assumed that the joint distribution of $[a_{ij}]$ was the same for micros and, separately, for macros. Assume a fixed group, therefore a fixed, single distribution for $[a_{im}]$.

We mentioned the irregularity of the observed data. Our approach was to impute values for patients at respective regular grids of points. For this, we employed a variation of natural cubic B-splines methods (20). For the jth subject, with measurements of $y_1, \ldots, y_n$ at times $t_1, \ldots, t_n$, write $y = (y_1, \ldots, y_n)'$ and $f(t) = [f(t_1), \ldots, f(t_n)]'$, where the prime denotes transpose; $y$ is random, with covariance matrix $V$. We minimized

$$\sum_{i=1}^{n} (y_i - \hat{f}(t_i))^2$$

for $\lambda$ that was chosen optimally by a method that will be described below. This has the effect of weighting observations equally. There was a common fitted value of $\lambda$ for micros, and also one for macros. Here, $t_1$ is the time of the first measurement ($y_1$) of GFR for the ith patient, and $t_n$ is the time of the last measurement (imputed if ESRF or dialysis); $c_1$ was taken to be just less than $t_1$, while $d_i$ was taken to be just larger than $t_n$. $\hat{f}$ is an estimate of $f$, again fit separately for micros and macros. The fits respected differences in $[\sigma_j^2]$. The fitted smoothing spline values were used to impute GFR at 6-mo intervals over the range of observation.

Assume that a group, micros or macros, is fixed. Write $\Sigma_n$ for the covariance matrix of $[a_{im}]$, with $\Sigma_{mm}$ its (m, m') element. Then, the covariance of $y_j(t_i)$ and $y_j'(t_{i'})$ is

$$\text{Cov}(y_j(t_i), y_j'(t_{i'}')) = \delta_{ij} \delta_{ij'} \sigma_j^2 + \sum_{m=1}^{M} B_m(t_i) B_m(t_{i'}) \Sigma_{mm} + \sum_{m=1}^{M} B_m(t_i) B_{m'}(t_{i'}) \Sigma_{mm}.$$  \hfill (A2)

Here, $\delta_{ij} = 1$ if $j = j'$, and 0 otherwise; similarly for $\delta_{ij'}$. The development thus far entails that for the $k$th subject

$$y_i = X_i a_i + \epsilon_i$$

where $X_i$ is a matrix of B-spline values at times observations were available for the $i$th subject. As before, $a_i$ and $\epsilon_i$ were taken to be independent, the covariance matrix of $a_i$ being $\Sigma_n$. The covariance matrix of errors of measurement $\epsilon_i$ is the matrix $D_i; D_i$ has equal elements for groups of $n_i$ observations. Thus the covariance matrix of $y_i$ is $X_i \Sigma_n X_i' + D_i$. For the $k$th subject, the approximation by smoothing splines and fitting by B-splines produced an estimated $\hat{a}_k$, call it $\hat{a}_k$, where for some matrix $M_k$, $\hat{a}_k = M_k y_k$.

The covariance matrix of $\hat{a}_k$ is then

$$(M_k X_k) \Sigma_n (M_k X_k)' + M_k D_k M_k'$$  \hfill (A3)

However, Rank ($M_k X_k) \leq n_i$. Even when $M_k X_k$ is not singular, it is usually very poorly conditioned. Either way, we could not use (A3) naively to solve for $\Sigma_n$.

We began by estimating $\Sigma_n$ with the naive estimate $\hat{\Sigma}_n$. The empirical covariance matrix of simple estimators $\hat{a}_k$. The principle axis theorem for quadratic forms enables us to write $\hat{\Sigma}_n$ as $\Sigma_n = E A \Sigma_n A'E$, where $A_1$ is real and diagonal and is unique up to permutations of its diagonal elements; E is orthogonal. Because the coordinates of $\hat{a}_k$ come from distinct observed values, it is not clear a priori that all eigenvalues of $A_1$ are positive. The “true” eigenvalues are necessarily nonnegative. Therefore, we set any negative eigenvalues to 0, creating the real, diagonal, nonnegative matrix $\hat{A}_1$. Write $\hat{\Sigma}_n = E \hat{A}_1 E'$. There are decision-theoretic and practical reasons for shrinking $\hat{\Sigma}_n$ toward a suitable multiple of the identity. See, for example, Ref. 8 and its references.

We write

$$\hat{\Sigma}_{final} = \hat{\Sigma}_n + (1-\theta)\epsilon_{med} I$$  \hfill (A4)

where $0 < \theta < 1; \epsilon_{med}$ is the median eigenvalue of $\Lambda$ (invariably positive); and I is the identity matrix of the correct dimension. The process of cross-validating is thus jointly over ($\lambda, \theta$), with one optimal choice for micros and one for macros. The optimal choice for each group was determined by successively deleting all observations for a fixed individual and point in time and using the deleted observations to assess the quality of fit. This “best” estimate $\hat{\Sigma}_{final}$ entails an estimate of $V$, which then figures in $A_2$; and the process begins again. In view of $A_2, A_3,$ and $A_4$, this algorithm is seen to be iterative and extremely computer intensive. For further details, see Ref. 3.

The process for fitting GFR curves thus far may seem excessively elaborate to readers if the conventional approach of modeling the evolution of GFR over time by a simple straight line is adequate. We studied by simulation the question of whether fitting by other than straight lines was necessary. To begin, we simulated data by fitting straight lines through the raw data, and then added simulated random errors, mean 0, normal errors using faithful standard deviations at each time point, producing 100 simulated data sets for each person, micro or macro. It was not feasible computationally to rerun our procedure completely on the simulated data, so we ran a procedure for which we assumed the estimates of the smoothing ($\lambda$) and shrinkage ($\theta$) parameters fitted for each of the two groups that comprise the original data. There remained the issue of inferring the covariance structure of the spline coefficients $[a_{im}]$. For each simulated data set, given the above assumption, we iterated the computation of the covariance three times. We calculated the difference between residual sums of squared errors
for the linear and nonlinear fits, and checked for how many of the sets of simulated values this difference was more extreme than it was for the real data. Our finding was that for 9 of the 26 micros and 5 of the 22 macros, this simulated statistic was “significant” at a level <0.05. To us, the data seemed to require spline fits that were not affine functions of time, so we recommend our approach for all patients.

Statistical comparisons of the GFR courses (progressors vs. non-progressors over 8 yr and the postprogresors course of initially microalbuminuric subjects vs. subjects screened as macroalbuminuric over 4 yr) were made on five-component vectors. The components were five inferred GFR values evenly spaced over the duration of follow-up for each subject (i.e., either 4 or 8 yr). To be precise, fits to $a_{on}$ in $A_f$ entail fits of $y_{ijk}$, thus of fits to the five-dimensional vectors. Two sets of five-dimensional vectors were compared by a two-sample Hotelling’s $T^2$ statistic (15). Rather than trust the (approximate) null distribution of $T^2$, we repeated the fitting for the common parameters should therefore be conservative.

For the comparison of the micro progressors and nonprogressors, we attempted the same procedure and found that for many of the permutations we were unable to calculate fits because of singular estimates of relevant covariance matrices. If subjects were homogeneous by covering diagnosis in terms of the probability mechanisms that generated them, as holds under the null hypothesis, then the null hypothesis of “no difference” is false, then having suboptimal fitting should render the permutation test we performed conservative.

For the comparison of the micro progressors and nonprogressors, we attempted the same procedure and found that for many of the permutations we were unable to calculate fits because of singular estimates of relevant covariance matrices. If subjects were homogeneous by covering diagnosis in terms of the probability mechanisms that generated them, as holds under the null hypothesis, then the difference using separate parameters should be larger; a difference were found using common parameters, then the large number of simulations feasible, we did fewer iterations and a more coarse search for a difference between micro progressors and macros, we randomly permuted the group assignments 100 times and then refitted the two permuted groups separately, with each group having its own pair of parameters ($\lambda, \theta$) before calculating $T^2$. To make these computations feasible, we did fewer iterations and a more coarse search for the minimizations than were done when fitting the real data. However, to make the test fair, we also repeated the fitting for the original data with the same restrictions and calculated $T^2$ from these fits. Note that if the null hypothesis of “no difference” is then valid, having suboptimal fitting should render the permutation test we performed conservative.

For the comparison of the micro progressors and nonprogressors, we attempted the same procedure and found that for many of the permutations we were unable to calculate fits because of singular estimates of relevant covariance matrices. If subjects were homogeneous by covering diagnosis in terms of the probability mechanisms that generated them, as holds under the null hypothesis, then the difference using separate parameters should be larger; a difference were found using common parameters, then the large number of simulations feasible, we did fewer iterations and a more coarse search for the minimizations than were done when fitting the real data. However, to make the test fair, we also repeated the fitting for the original data with the same restrictions and calculated $T^2$ from these fits. Note that if the null hypothesis of “no difference” is then valid, having suboptimal fitting should render the permutation test we performed conservative.

For the comparison of the micro progressors and nonprogressors, we attempted the same procedure and found that for many of the permutations we were unable to calculate fits because of singular estimates of relevant covariance matrices. If subjects were homogeneous by covering diagnosis in terms of the probability mechanisms that generated them, as holds under the null hypothesis, then the difference using separate parameters should be larger; a difference were found using common parameters, then the large number of simulations feasible, we did fewer iterations and a more coarse search for the minimizations than were done when fitting the real data. However, to make the test fair, we also repeated the fitting for the original data with the same restrictions and calculated $T^2$ from these fits. Note that if the null hypothesis of “no difference” is then valid, having suboptimal fitting should render the permutation test we performed conservative.


