Published in final edited form as:

Stat Methods Med Res. 2017 June ; 26(3): 1519–1531. doi:10.1177/0962280215584109.

Testing the trajectory difference in a semi-parametric longitudinal model

Feiyang Niu,

Department of Statistics, University of Virginia, Charlottesville, VA 22904 USA

Jianhui Zhou,

Department of Statistics, University of Virginia, Charlottesville, VA 22904 USA

Thu H. Le, and

Division of Nephrology, Department of Medicine, University of Virginia, Charlottesville, VA 22908 USA

Jennie Z. Ma

Division of Biostatistics, Department of Public Health Sciences, University of Virginia, Charlottesville, VA 22908 USA

Abstract

Motivated by a genetic investigation on the progressive decline of renal function in a clinical trial study of kidney disease, we develop a practical test for evaluating the group difference in trajectories under a semi-parametric modeling framework. For the temporal patterns or trajectories of longitudinal data, B-splines are used to approximate the function non-parametrically. Such approximation asymptotically converts the problem of testing trajectory difference into the significance test of regression coefficients that can be simply estimated by generalized estimating equations. To select the optimal number of inner knots for B-splines, a cross-validation procedure is performed using the criterion of generalized residual sum of squares. The new proposed test successfully detects the significant difference of underlying genetic impact on the progression of renal disease, which is not captured by the parametric approach.

Keywords

B-spline; longitudinal data; generalized estimating equation; semi-parametric; trajectory testing

1 Introduction

With the development of modern medical statistics, longitudinal data have been extensively analyzed using parametric models, either marginal or mixed effects models, in many biomedical applications. These parametric approaches for longitudinal data allow us not only to evaluate the effects of population-level risk factors while accounting for subject-specific effects,¹ but also to model the temporal patterns and test for group difference in trajectories over time. For example, in studying patients with hypertensive kidney disease, researchers are interested in comparing the effects of antihypertensive drugs to prevent progressive decline of renal function, which leads to kidney failure. The objective is to test

the difference in the patterns of renal function decline over time.² In the parametric settings, response trajectory patterns are usually modeled parametrically as polynomials of time, and group difference in trajectory patterns is commonly tested by the Group-by-Time interaction.³

While parametric models are highly useful, they require specific form for the mean function of response over time, which can be challenging in practice.⁴ Moreover, questions often arise about the adequacy of the model assumption and the potential impact of model misspecification on the analysis.⁵ Alternatively, non-parametric or semi-parametric regression methods for longitudinal data offer more flexible modeling options by relaxing the restrictions of parametric models on the temporal trend. There are abundant examples of application of semi-parametric regression that is based on penalized regression splines and mixed models and examples of likelihood ratio test of curve difference for cross-sectional data in Ruppert et al.⁶ Particularly, the implementation of smoothing methods in standard software, such as R, has stimulated the applications of non-parametric and/or semi-parametric modeling to clinical studies.⁷⁻⁹ However, the direct implementation of non-parametric and/or semi-parametric modeling of longitudinal data does not render the test for group differences in trajectories over time. To our knowledge, ready-to-use statistical test is only available for testing the group differences at a specific time point, but not for testing the group trajectory differences, or longitudinal response profile differences. Lack of such ready-to-use test in trajectory analysis may result in limited broader application of non-parametric and semi-parametric methods in clinical studies.

In this paper, we study a semi-parametric model with non-parametric time effect on trajectory for longitudinal data and develop a statistical test that can ascertain the group trajectory differences. Our developed test can be readily applied to testing trajectory differences over the entire follow-up period, or over a subinterval of interest. Our modeling and statistical test development are motivated by a genetic investigation on the progressive decline in renal function in the African American Study of Kidney Disease and Hypertension (AASK). The AASK study was a multicenter randomized clinical trial to test the effectiveness of three antihypertensive medications and two levels of blood pressure (BP) control on the progression of hypertensive kidney disease in a 3×2 factorial design.¹⁰ It included 1094 African Americans aged 18 to 70 years old with hypertensive kidney disease, measured by glomerular filtration rates (GFR) of 20 to 65 mL/min per 1.73m². During February 1995 and September 1998, patients were randomly assigned to one of two mean arterial pressure (MAP) goals, 102 to 107 mm-Hg or 92 mm-Hg, and to initiate treatment with an angiotensin-converting enzyme (ACE) inhibitor, a β -blocker, or a dihydropyridine calcium channel blocker. Details on the original AASK study design and conduct are available elsewhere.^{2,10,11} The objective of our genetic study is to evaluate the association of Glutathione S-transferase γ 1 (*GSTM1*) gene with the kidney function, using the data available in the AASK Study. Specifically, the researchers are interested in whether patients with the *GSTM1*-null allele would have a more accelerated progression of kidney disease than those with *GSTM1*-active variant.¹²

2 Methodology

In a general longitudinal study, suppose that there are m subjects, n_i observations are collected at time $\{0 \leq t_{ij}\}_{j=1}^{n_i}$ for the i th subject, and we have $n = \sum_{i=1}^m n_i$ observations in total. Let Y_{ij} and $\mathbf{Z}_{ij} = (z_{ij1}, z_{ij2}, \dots, z_{ijq})^T$ denote respectively the measurement of response and the q covariates, such as age, gender etc., for the i th subject at time t_{ij} , where $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n_i$. Suppose subjects are distributed over K groups, and the interest is to test the differences in the longitudinal patterns among these K groups. To accurately characterize the longitudinal patterns and avoid model mis-specification, we consider the time effect non-parametrically in the following semi-parametric model,

$$Y_{ij} = \mathbf{Z}_{ij}^T \boldsymbol{\alpha} + f_0(t_{ij}) + \sum_{k=1}^{K-1} \mathbf{I}(g_i = k) \delta_k(t_{ij}) + \epsilon_{ij}, \quad (1)$$

for $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n_i$, where $\boldsymbol{\alpha}$ is a q -dimensional coefficient vector corresponding to the q covariates; $\mathbf{I}(g_i = k)$ is the indicator function for group g_i , which is a categorical variable indicating the group membership that the i th subject belongs to; f_0 is an unspecified smooth function of time depicting the trajectory of response for subjects in Group K (as reference group), and δ_k 's are the trajectory differences in response between the reference group and the k th group, $k = 1, 2, \dots, K-1$; and ϵ_{ij} is the random error. We assume that the maximum follow-up lengths are equal among K groups, which is denoted as T . The semi-parametric model (1) has been studied in Lin and Carroll,¹³ He *et al.*,¹⁴ and Leng *et al.*¹⁵ among others that mainly focuses on the inference of the parametric part. Here, we focus on the inference of the nonparametric part. By taking the advantage of the local property of the B-spline basis functions, our method has the flexibility of testing the trajectory difference on the overall time interval or a specific subinterval. Note that both f_0 and δ_k 's are adjusted for the covariates $\mathbf{Z}_{ij}$. Thus, the trajectory differences between the k th group and the reference group K can then be detected by testing whether the functions $\delta_k(t)$ are zero, i.e.,

$$\delta_k(t) = 0, \quad (2)$$

for any $t \in [0, T]$ and $k = 1, 2, \dots, K-1$.

We approximate the smooth functions f_0 and δ_k 's by B-splines.¹⁶ Given the spline order r and a partition $0 = t_0 < t_1 < \dots < t_s < t_{s+1} = T$, where s is the number of interior knots, we denote the normalized B-spline basis functions by $\{B_l\}_{l=1}^L$ for $L = s + r$. The functions f_0 and δ_k are then approximated by

$$\begin{aligned} f_0(t_{ij}) &= \sum_{l=1}^L \beta_{0l} B_l(t_{ij}) + e_0(t_{ij}) \\ \delta_k(t_{ij}) &= \sum_{l=1}^L \beta_{kl} B_l(t_{ij}) + e_k(t_{ij}), \end{aligned} \quad (3)$$

for $k = 1, 2, \dots, K-1$, where $e_0(t)$ and $e_k(t)$'s are approximation error functions that converge to zero uniformly at a certain rate with the increasing number of knots and sample size. In practice, low order of basis functions and equally spaced quantile inner knots are often used. Advantages of B-splines mainly attributes to its efficient computation and accurate approximation with a small number of knots.¹⁷ For simplicity, we use the same set of B-spline basis functions to approximate f_0 and δ_k . Combining (1) and (3) together, we have the model

$$Y_{ij} = \mathbf{Z}_{ij}^T \boldsymbol{\alpha} + \sum_{l=1}^L \beta_{0l} \mathbf{B}_l(t_{ij}) + \sum_{k=1}^{K-1} \sum_{l=1}^L \mathbf{I}(g_i=k) \beta_{kl} \mathbf{B}_l(t_{ij}) + \epsilon'_{ij}, \quad (4)$$

where $\epsilon'_{ij} = \epsilon_{ij} + e_0(t_{ij}) + \sum_{k=1}^{K-1} \mathbf{I}(g_i=k) e_k(t_{ij})$, converging to ϵ_{ij} uniformly.

2.1 Estimation and hypothesis testing

Let $\boldsymbol{\beta}_0 = (\beta_{01}, \beta_{02}, \dots, \beta_{0L})^T$, $\boldsymbol{\beta}_k = (\beta_{k1}, \beta_{k2}, \dots, \beta_{kL})^T$, for $k = 1, 2, \dots, K-1$, and $\epsilon'_i = (\epsilon'_{i1}, \epsilon'_{i2}, \dots, \epsilon'_{in_i})$, and denote

$$\mathbf{Z}_i = \begin{pmatrix} \mathbf{Z}_{i1}^T \\ \mathbf{Z}_{i2}^T \\ \vdots \\ \mathbf{Z}_{in_i}^T \end{pmatrix}_{n_i \times q} \quad \text{and} \quad \mathbf{X}_i = \begin{pmatrix} B_1(t_{i1}) & B_2(t_{i1}) & \cdots & B_L(t_{i1}) \\ B_1(t_{i2}) & B_2(t_{i2}) & \cdots & B_L(t_{i2}) \\ \vdots & \vdots & \ddots & \vdots \\ B_1(t_{in_i}) & B_2(t_{in_i}) & \cdots & B_L(t_{in_i}) \end{pmatrix}_{n_i \times L}.$$

Model (4) can thus be rewritten as

$$Y_i = \mathbf{Z}_i \boldsymbol{\alpha} + \mathbf{X}_i \boldsymbol{\beta}_0 + \sum_{k=1}^{K-1} \mathbf{I}(g_i=k) \mathbf{X}_i \boldsymbol{\beta}_k + \epsilon'_i, \quad (5)$$

for $i = 1, 2, \dots, m$, where $\mathbf{Y}_i = (Y_{i1}, Y_{i2}, \dots, Y_{in_i})^T$ are the repeated measurements of response on the i th subject. Then the generalized estimation equations (GEE) method, developed by Liang and Zeger,¹⁸ can be used to estimate the parameters in (5). The application of B-spline within GEE framework can be found in Hua and Zhang¹⁹ and He et

al.¹⁴ For the simplicity of notation, we let $\boldsymbol{\beta} = (\boldsymbol{\alpha}^T, \boldsymbol{\beta}_0^T, \boldsymbol{\beta}_1^T, \dots, \boldsymbol{\beta}_{K-1}^T)^T$ and

$\mathbf{U}_i = (\mathbf{Z}_i, \mathbf{X}_i, \mathbf{I}(g_i=1) \mathbf{X}_i, \dots, \mathbf{I}(g_i=K-1) \mathbf{X}_i)$. The GEE estimator

$\hat{\boldsymbol{\beta}} = (\hat{\boldsymbol{\alpha}}^T, \hat{\boldsymbol{\beta}}_1^T, \dots, \hat{\boldsymbol{\beta}}_{K-1}^T)^T$ is obtained by solving the following equations

$$\sum_{i=1}^m \mathbf{U}_i^T \boldsymbol{\Sigma}_i^{-1} (\mathbf{Y}_i - \mathbf{U}_i \boldsymbol{\beta}) = 0, \quad (6)$$

where $\Sigma_i = A_i^{1/2} R A_i^{1/2}$ is the covariance matrix of Y_i , A_i is the diagonal marginal variance of Y_i , and R is the working correlation matrix. The asymptotic variance of $\hat{\beta}$ is estimated by the sandwich formula in the GEE framework.

By approximating the functions $f_0(t)$ and $\delta_k(t)$ in Model (1) through B-splines, we essentially convert the problem of testing the group differences on the time effect, i.e., trajectories, over the entire follow-up period into the following hypothesis testing

$$H_0: \beta_k = \mathbf{0} \quad \text{versus} \quad H_1: \beta_k \neq \mathbf{0}, \quad (7)$$

for $k = 1, 2, \dots, K-1$. Given the asymptotic distribution of $\hat{\beta}$, the vector $\hat{\beta}_k$ asymptotically has the following distribution

$$\hat{\beta}_k \sim N(\beta_k, \hat{\Omega}_k),$$

where $\hat{\Omega}_k$ is the submatrix of $\hat{\Omega}$, the estimated asymptotical covariance matrix of $\hat{\beta}$ by the sandwich formula of GEE. A Wald-type test statistic,

$$W^2 = \hat{\beta}_k' \hat{\Omega}_k^{-1} \hat{\beta}_k, \quad (8)$$

has an asymptotical χ^2 distribution with L degrees of freedom (df), where L is the length of $\hat{\beta}_k$.

Besides testing the group differences for the overall trajectories, another advantage of our developed method is its ability to test the group differences of the time effect or trajectories over a subinterval of the follow-up period. This ability attributes to the locality property of the B-spline basis functions. Indeed, once the inner knots are allocated, the function estimate for δ_k is completely determined by a consecutive subset of parameters in $\hat{\beta}_k$ that are associated with the subinterval. That is, this subset of parameters is uniquely identified according to the locality property. Therefore, the group differences over the subinterval of interest can be detected by testing whether the identified subset of parameters in β_k are simultaneously zero as in (8). An example is shown in Section 3.2 below.

2.2 Knots selection

To implement our developed method in Section 2.1, a key practical issue is to specify data-adaptive B-spline basis functions. In non-parametric regression, low order of spline is usually preferred, such as linear, quadratic, or cubic spline with the spline order r taking value of two, three, or four respectively. In our study, quadratic B-spline basis functions (i.e., $r = 3$) are used in order to keep the model sufficiently flexible yet less complicated, and inner knots are selected as equally spaced quantile knots as frequently suggested in the literature. The number of inner quantile knots s is selected by 10-fold cross validation using

the generalized residual sum of squares (GRSS) as the selection criterion (see Section 3 for an example). The GRSS is defined as

$$GRSS(s) = \sum_{i=1}^m (Y_i - \hat{Y}_i)' \hat{\Sigma}_i^{-1} (Y_i - \hat{Y}_i), \quad (9)$$

where $\hat{Y}_i = U_i \hat{\beta}$, $i = 1, 2, \dots, m$, are fitted values for the i th subject and $\hat{\Sigma}_i$ is the estimated working covariance matrix based on the GEE estimates. The GRSS criterion balances both bias and variation in estimating the functions $f_0(t_{ij})$ and $\delta_k(t_{ij})$ using B-spline approximation. The final selection for number of inner knots will be the model that minimizes the GRSS criterion.

3 Analysis of real data

3.1 Background

As described in Section 1, 1094 African American patients with chronic kidney disease were enrolled into the AASK study between 1995 and 1998. Most of them were followed up to September 2001, and those on the calcium channel blocker arm were terminated in September 2000 based on the recommendation of the data and safety monitoring board.^{2,10} Beyond the baseline measurements such as age and gender, the longitudinal kidney function, measured by GFR in mL/min per 1.73m², was assessed at baseline, three, six, and every three months thereafter. Approximately 850 patients who remained available in 2002 further consented for the AASK Genomics Study. A total of 692 patients with valid DNA samples were successfully genotyped and included in our analysis. Our genetic study aimed to evaluate the association of *GSTM1* gene with the progression of the kidney disease. Earlier studies have demonstrated that patients who carry *GSTM1* null allele, *GSTM1(0)*, have increased risk of cardiovascular disease, and this is thought to be due to the reduced ability or inability to handle oxidative stress and the resultant cellular damage.

In this study, we would like to evaluate whether patients with the *GSTM1(0)* allele have more accelerated GFR decline than those with *GSTM1*-active variant. Thus we are interested in testing whether the GFR declining patterns or trajectories over time are significantly different between patients with *GSTM1(0)* and those without. Detailed patient characteristics for the 692 patients in the study cohort by *GSTM1* genotype have been reported previously (Table 1 in Chang et al.¹²). Briefly, about 27% patients were classified in *GSTM1(0)* group and 73% in *GSTM1*-active group. The mean age for the study cohort was 54.2 ± 10.6 years, 59% were male, and 51% had history of cardiovascular disease. Patients were distributed evenly in the randomization factors (blood pressure control level and antihypertensive drug group). These baseline characteristics were included in our semi-parametric model.

3.2 Results

To capture the true decline of renal function in our analysis, we model the change in GFR from baseline measure instead of actual GFR. To adjust for potential confounding factors,

patient baseline characteristics such as age, gender, MAP, history of cardiovascular disease, blood pressure control level and antihypertensive drug group are included as the covariates $\mathbf{Z}_{ij}$ in model (1). The first-order autoregressive correlation structure (AR(1)) is assumed as the working correlation matrix for GEE among the repeated measurements of GFR response within the same subject. To estimate the parameters, $\alpha, \beta_0, \dots, \beta_{K-1}$, in model (4), we utilize the data up to 60 months, where sufficient patients with GFR measurements are observed. Thus the overall follow-up period is from randomization (time 0) to 60 months. Let s denote the number of inner knots placed within $[0, 60]$, we select its value based on the best model fitting by 10-fold cross validation with the GRSS criterion. Specifically, we let s vary from 0 to 10. The value of GRSS for the model with s inner knots is calculated as in (9) and shown in Figure 1. It appears that the best model is the one with $s = 1$ optimal inner knot, making it as median of all visited time points across subjects.

Therefore, in the study of genetic effect of *GSTM1*, we use quadratic B-spline basis function with one inner knot to fit the longitudinal change in GFR. Figure 2 shows the observed and fitted time effects or trajectories for the two genetic groups, adjusted for patient baseline characteristics. Two fitted functions for the change in GFR from baseline clearly show decreasing trends, indicating deterioration of kidney function over the entire follow-up period. Also, for the patients in both genetic groups, there are apparent curvatures in the trajectories for the change in GFR. Such non-linear trajectory patterns demonstrate the advantage of semi-parametric modeling approach, which would not be captured easily in a parametric model. In addition, the trajectories for the change in GFR between the two groups apparently decline at different rates. The *GSTM1(0)* group declines slower in the first 22 months, crossing with the *GSTM1*-active group approximately at 22 months and then accelerates afterwards.

To test the trajectory difference between the two groups over the follow-up period, we apply the hypothesis testing under semi-parametric framework as proposed in Section 2. The results in Table 1 show that the *GSTM1(0)* group is marginally different in GFR declining pattern from the *GSTM1*-active group over $[0, 60]$ months, with p-value 0.0753. Meanwhile, the following six commonly used parametric models, i.e. linear marginal/mixed effects models, piecewise linear marginal/mixed effects models, and quadratic marginal/mixed effects models, have also been applied to the data. The linear marginal model depicts GFR change trajectory by linear form of time and linear mixed effects model additionally includes random intercept and random time slope with both models adjusted for the same set of baseline characteristic covariates as in semi-parametric model. The piecewise linear marginal/mixed effects models and the quadratic marginal/mixed effects models are specified in a similar fashion. The p-values are obtained by testing the significance of interaction terms between genotype group and time. None of these parametric models detected significant GFR trajectory difference between *GSTM1(0)* group and *GSTM1*-active group, as summarized in Table 1.

The fitted changes in GFR by our semi-parametric model clearly reveal different declining patterns of kidney progression between the two groups after 20 months. We further test if the two trajectories are significantly different after 20 months, the long term effect. As the optimal inner knot is at 20 months, the subinterval from 20 to 60 months corresponds to the

B-spline coefficients β_{12} , β_{13} and β_{14} in model (4) by the locality property as discussed in Section 2.1. The testing results in Table 2 show that the trajectory of change in GFR for the *GSTM1(0)* group is indeed significantly different from that in the *GSTM1*-active group for the long-term follow-up period (p-value=0.038).

Model diagnostics is an important aspect in data analysis for evaluating model goodness-of-fit, for detecting outliers, and for identifying influential observations. In the absence of likelihood ratio tests, model diagnostic tools are needed for the GEE approach where correlated data generally arise.²⁰ By extending Akaike's information criterion to the GEE framework, Pan²¹ proposed to use the quasi-likelihood constructed under the working independence model with the naive and robust covariance estimates of estimated regression coefficients. In our study, we applied Pan's quasi-likelihood under the independence model criterion (QIC) to our semi-parametric model as well as to the six parametric models. Our semi-parametric model achieved the minimum QIC as shown in Table 1. This further confirms the better goodness-of-fit of our semi-parametric model over those parametric models.

In our semi-parametric model, some baseline characteristics have significant effects on the response, as presented in Table 3. Specifically, older age and higher baseline GFR are apparently associated with greater decline of GFR. Compared with ACE inhibitor, patients on β -blocker have a greater decline in GFR, while those on calcium channel blocker actually have an increase in the change of GFR. Our results on these covariate effects are consistent with previous reports.^{12,22}

To further explore the complexity of the study, we also investigated GFR trajectory difference among genotype $\times$ treatment groups and among genotype $\times$ blood pressure goal groups. Patients in *GSTM1(0)* and β -blocker and those in *GSTM1(0)* and low blood pressure goal are, respectively, considered as reference group in each of the two analysis. Fitted GFR change curves are presented in Figure 3. Hypothesis testing results on progression of GFR change are summarized in Table 4. This result shows that GFR trajectories decline at variable rates with different shapes for the combinations of genotype with antihypertensive drug or for the combinations of genotype with blood pressure goal. Indeed, patients in the *GSTM1(0)* $\times$ β -blocker group (the reference group) had the steepest GFR decline. The GFR trajectories in calcium channel blocker group at either *GSTM1* level as well as in the *GSTM1*-active $\times$ ACE inhibitor group appear to decline slower and they are significantly different from that of the reference group. Similarly, patients with *GSTM1*-active genotype apparently have slower deterioration in kidney function, and their trajectories are significantly different from that in the *GSTM1(0)* group. Although the trajectories for patients with *GSTM1(0)* seemingly differ for the two blood pressure goals, the difference is not statistically significant.

3.3 Simulation

Our proposed test for trajectory difference relies on the Wald-type test statistics defined in (8). In this section, we conduct a simulation study to investigate the finite sample performance of the proposed test based on AASK study data. Specifically, a total number of

500 patients are generated with the same set of baseline covariates included in our semi-parametric model; i.e., for the i^{th} patient,

$$y_{ij} = \mathbf{Z}_{ij}^T \boldsymbol{\alpha} + f_0(t_{ij}) + f_1(t_{ij}) \mathbf{I}(g_i = \text{GSTM1-active}) + \epsilon_{ij}, \quad (10)$$

for $i = 1, 2, \dots, 500$ and $j = 1, 2, \dots, n_i$ where n_i is a random integer from 6 to 12; $\mathbf{Z}_{ij}$ is a 7-dimensional baseline characteristic covariates vector including age, gender, MAP, baseline GFR, history of cardiovascular disease, blood pressure control level and antihypertensive drug group; $\boldsymbol{\alpha}$ is the coefficient vector corresponding to the 7 baseline covariates and is specified to be the same as the estimated one from the real data in Section 3.2; $f_0(\cdot)$ is the underlying true GFR change curve for subjects in for *GSTM1(0)* group and $f_1(\cdot)$ is the trajectory differences in GFR change between *GSTM1(0)* and *GSTM1-active* groups, both of which are specified be similar to the two estimated GFR change curves in Section 3.2; t_{ij} 's are uniformly generated from 0 to $T = 60$; $\mathbf{I}(\cdot)$ is the indicator function indicating the genetic group membership of the i^{th} subject; $\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \epsilon_{i2}, \dots, \epsilon_{in_i})^T$ is the random vector having AR(1) correlation structure with $\sigma = 13.3$ and $\rho = 0.86$, same as those estimated from the data.

Our proposed semi-parametric model is applied to the simulated data with three different working correlation structures, AR(1) (the true structure), compound symmetry (CS), and independence. The testing for trajectory difference in the GFR change is based on the Wald test with 5% significance level. Out of the $M = 500$ simulation runs, power of the test is reported as the percentage of successfully detecting the curve difference. Furthermore, the goodness-of-fit of estimation is evaluated by the weighted average squared error (WASE) defined as

$$WASE = \frac{1}{N} \sum_{i=1}^n \sum_{j=1}^{n_i} \frac{(f_k(t_{ij}) - \hat{f}_k(t_{ij}))^2}{\text{range}(f_k(t))}, \quad (11)$$

where $N = \sum_{i=1}^n n_i$ and $\text{range}(f_k(t)) = \max_{t \in [0, T]} f_k(t) - \min_{t \in [0, T]} f_k(t)$. Both testing power and WASE are summarized in Table 5. In addition, boxplots of WASE are provided in Figure 4.

Out of the 500 runs, 94.6% of the runs successfully detected the GFR change trajectory difference between the two genetic groups for AR(1) working correlation, 91.8% for compound symmetry, and 85.8% for working independence. Numerical outcomes in Table 5 show that accounting for the within cluster correlation improves both testing results and estimation precision.

4 Discussion

In this work, we use a semi-parametric model to capture the non-linear decline of longitudinal renal function and test the trajectory difference over the follow-up period between the two genotype groups. While the existing methods can only test the difference at

an individual time point, our interests are focused on testing the functional effects over the entire or subinterval of follow-up period. By utilizing B-spline approximation, we effectively convert the problem of testing trajectory differences over time into a linear model for repeated measurements. Thus our developed testing method can be easily implemented through testing the selected parameters estimated by GEE, using existing commonly used statistical packages. Our developed method has the flexibility of testing the trajectory differences over the entire follow-up period, or a subinterval of interest. In addition, the semi-parametric model can accommodate various trajectory patterns, which is particularly attractive and important in flexible model fitting, and thus avoid the potential pattern misspecification in parametric model setting. However, to estimate each trajectory function reliably over the testing period, the B-spline approximation method requires that the maximum follow-up lengths are equal among groups under comparison and patient visit time points to scatter evenly within each group. Thus, our developed method is more efficient when the patient visit time is less skewed, though the longitudinal data are not required to be balanced.

In our analysis, we first selected the number of knots for the B-splines, and the trajectory differences were tested using the model with the selected knots. Therefore, our analysis did not account for the uncertainty introduced by model selection. Hjort and Claeskens²³ studied the post-selection inference issues. Following the data splitting idea in Wasserman and Roeder²⁴ and Meinshausen, Meier, and Bühlmann²⁵ to address this issue, we have explored to select the knots using a random sample of 30% of the subjects, and estimate the selected model for inferences using the remaining 70% subjects. The p-values for overall and long-term effects of *GSTM1* gene are 0.0783 and 0.0388, respectively, which are very close to the results in Table 2 using the full data set.

With our developed method, it is more likely to detect the true underlying trajectory differences that could be missed in parametric modeling, as shown in our clinical example. Our method can also be used to test the trajectory interaction between two or more factors by considering combined levels of these factors as shown in Figure 3 and Table 4. In addition, we plan to test the interaction effect of *GSTM1* and *APOL1* risk variants, as a recent study found that African Americans with the *APOL1* risk variants experience faster progression of chronic kidney disease and have a significantly increased risk of kidney failure.²⁶

Acknowledgement

We are most grateful to the two referees and the editor for their constructive comments and suggestions. We thank the AASK trial participants and Ancillary Studies Committee for granting access to the DNA and Trial data. This work was supported in part by the National Institutes of Health Grant R01 DK094907 to T. H. Le.

References

1. Fitzmaurice, GM., Laird, NM., Ware, JH. Applied Longitudinal Analysis. John Wiley and Sons; New York: 2004.
2. Wright JT Jr, Bakris G, Greene T, et al. Effect of blood pressure lowering and antihypertensive drug class on progression of hypertensive kidney disease: results from the AASK trial. Journal of the American Medical Association. 2002; 288:2421–2431. [PubMed: 12435255]

3. Singer JD. Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics*. 1998; 23:323–355.
4. Lin DY, Ying Z. Semiparametric and nonparametric regression analysis of longitudinal data. *Journal of the American Statistical Association*. 2001; 96:103–113.
5. Hoover DR, Rice JA, Wu CO, Yang LP. Nonparametric smoothing estimates of time-varying coefficient models with longitudinal data. *Biometrika*. 1998; 85:809–822.
6. Ruppert, D., Wand, MP., Carroll, RJ. *Semiparametric Regression*. Cambridge University Press; Cambridge: 2003.
7. Ngo L, Wand MP. Smoothing with Mixed Model Software. *Journal of Statistical Software*. 2004; 9:1–54.
8. Durban M, Harezlak J, Wand MP, Carroll RJ. Simple fitting of subject-specific curves for longitudinal data. *Statistics in Medicine*. 2005; 24:1153–1167. [PubMed: 15568201]
9. Chen J, Johnson BA, Wang XQ, O'Quigley J, Isaac M, Zhang D, Liu L. Trajectory analyses in alcohol treatment research. *Alcohol Clin Exp Res*. 2012; 36:1442–1448. [PubMed: 22525000]
10. Gassman JJ, Greene T, Wright JT Jr, et al. Design and statistical aspects of the African American Study of Kidney Disease and Hypertension (AASK). *Journal of the American Society of Nephrology*. 2003; 14:S154–S165. [PubMed: 12819322]
11. Agodoa LY, Appel L, Bakris GL, et al. African American Study of Kidney Disease and Hypertension (AASK) Study Group. Effect of ramipril vs amlodipine on renal outcomes in hypertensive nephrosclerosis: a randomized controlled trial. *Journal of the American Medical Association*. 2001; 285:2719–2728. [PubMed: 11386927]
12. Chang J, Ma JZ, Zeng Q, et al. Loss of GSTM1, a NRF2 target, is associated with accelerated progression of Hypertensive Kidney Disease in the African American Study of Kidney Disease (AASK). *American Journal of Physiology - Renal Physiology*. 2013; 304:F348–F355. [PubMed: 23220723]
13. Lin X, Carroll RJ. Semiparametric regression for clustered data using generalized estimation equations. *Journal of the American Statistical Association*. 2001; 96:1045–1056.
14. He X, Fung WK, Zhu Z. Robust estimation in generalized partial linear models for clustered data. *Journal of the American Statistical Association*. 2005; 100:1176–1184.
15. Leng C, Zhang W, Pan J. Semiparametric mean-covariance regression analysis for longitudinal data. *Journal of the American Statistical Association*. 2010; 105:181–193.
16. Schumaker, LL. *Spline Functions: Basic Theory*. Wiley-Interscience; New York: 1981.
17. He X, Shen L. Linear regression after spline transformation. *Biometrika*. 1997; 84:474–481.
18. Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika*. 1986; 73:13–22.
19. Hua L, Zhang Y. Spline-based semiparametric projected generalized estimating equation method for panel count data. *Biostatistics*. 2012; 13:440–454. [PubMed: 21926062]
20. Oh S, Carriere KC, Park T. Model diagnostic plots for repeated measures data using the generalized estimating equations approach. *Computational Statistics & Data Analysis*. 2008; 53:222–232.
21. Pan W. Akaike's information criterion in generalized estimating equations. *Biometrics*. 2001; 57:120–125. [PubMed: 11252586]
22. Rowe JW, Andres R, Tobin JD, Norris AM, Shock NW. The effect of age on creatinine clearance in men: a cross-sectional and longitudinal study. *The Journals of Gerontology*. 1976; 31:155–63.
23. Hjort NL, Claeskens G. Frequentist model average estimators. *Journal of the American Statistical Association*. 2003; 98:879–899.
24. Wasserman L, Roeder K. High dimensional variable selection. *The Annals of Statistics*. 2009; 36:1567–1594.
25. Meinshausen N, Meier L, Bühlmann P. P-values for high-dimensional regression. *Journal of the American Statistical Association*. 2009; 104:1671–1681.
26. Parsa A, Kao WHL, Xie D, et al. APOL1 risk variants, race, and progression of chronic kidney disease. *The New England Journal of Medicine*. 2013; 369:2183–2196. [PubMed: 24206458]

Figure 1.

Cross-validation for selecting number of inner knots s . The y -axis represents the model selection criterion of GRSS, and the x -axis is the number of inner knots placed within $[0, 60]$ months. The solid red point indicates where the minimum GRSS is attained.

Figure 2. Observed and fitted curves of changes in GFR from baseline in *GSTM1(0)* Group (black solid circles for observed and black solid line for fitted) and *GSTM1*-active Group (red hollow circles for observed and red dashed line for fitted). Each circle represents GFR change averaged on that particular scheduled time point. The y -axis represents the fitted change in GFR from baseline, and the x -axis is the follow-up time in months.

(a) Genotype $\times$ Treatment(b) Genotype $\times$ Blood pressure goal**Figure 3.**

(a): Fitted curves of changes in GFR among Genotype (solid line for *GSTM1(0)* Group and dashed line for *GSTM1-active* Group) $\times$ Treatment (red for ACE inhibitor, black for β -blocker and green for calcium channel blocker); (b): Fitted curves of changes in GFR among Genotype (solid line for *GSTM1(0)* Group and dashed line for *GSTM1-active* Group) $\times$ Blood pressure goal (black for low BP goal and red for usual BP goal).

Figure 4. Boxplot of WASE of fitted semi-parametric model with three different working correlation structure of $M = 500$ simulation runs

Table 1

Model comparison.

	P-value ¹	QIC ²	N. of params ³
Semi-parametric model (1)	0.0753	28088	18
Marginal models			
Linear model	0.1307	28095	12
Piecewise linear model (breakpoint at 20 months)	0.1691	28098	14
Quadratic model	0.2102	28104	14
Mixed effects models			
Linear model	0.1701	28102	13
Piecewise linear model (breakpoint at 20 months)	0.3126	28114	15
Quadratic model	0.3649	28120	15

¹P-value for trajectory difference between two genetic groups²QIC: Pan's quasi-likelihood under the independence model criterion³N. of params: Number of parameters in model

Table 2

Hypothesis testing for genetic impact on progression of GFR.

	<i>GSTMI(0)</i> Group vs <i>GSTMI</i> -active Group		
	χ^2 statistic	<i>df</i>	P-value
Trajectory test for overall effect ([0, 60] months)	8.484	4	0.0753
Trajectory test for long-term effect ([20, 60] months)	8.425	3	0.0380

Table 3

Estimation results for factors adjusted in semi-parametric model.

Factor	Estimate	Standard error	P value
Age	−0.0851	0.0292	0.0036
Gender (Male vs. Female)	0.5213	0.6452	0.4190
Baseline GFR	−0.1048	0.0208	<0.0001
Baseline MAP	−0.0328	0.0183	0.0726
History of cardiovascular disease	0.2930	0.6065	0.6291
Drug group (β -blocker vs. ACE inhibitor)	−1.4368	0.6507	0.0272
Drug group (calcium channel block vs. ACE inhibitor)	2.0190	0.8899	0.0233
Blood pressure goal (102-107 vs. 92 mm-Hg)	−0.1947	0.5986	0.7450

Table 4Hypothesis testing for genotype $\times$ treatment and genotype $\times$ blood pressure goal on progression of GFR

	Overall trajectory test([0, 60] months)			Long-term trajectory test([20, 60] months)		
	χ^2 statistic	df	P-value	χ^2 statistic	df	P-value
Genotype $\times$ Antihypertensive Drug						
Reference group: <i>GSTM1</i> (0) $\times$ β -blocker						
<i>GSTM1</i> (0) $\times$ ACE inhibitor	4.502	4	0.3423	4.493	3	0.2129
<i>GSTM1</i> (0) $\times$ calcium channel blocker	25.156	4	<0.0001	24.840	3	<0.0001
<i>GSTM1</i> -active $\times$ ACE inhibitor	9.324	4	0.0535	9.107	3	0.0279
<i>GSTM1</i> -active $\times$ β -blocker	0.911	4	0.9230	0.905	3	0.8244
<i>GSTM1</i> -active $\times$ calcium channel blocker	12.140	4	0.0163	11.768	3	0.0082
Genotype $\times$ Blood pressure goal						
Reference group: <i>GSTM1</i> (0) $\times$ Low BP goal						
<i>GSTM1</i> (0) $\times$ Usual BP goal	9.473	4	0.0503	6.528	3	0.0886
<i>GSTM1</i> -active $\times$ Low BP goal	8.641	4	0.0707	8.537	3	0.0361
<i>GSTM1</i> -active $\times$ Usual BP goal	7.954	4	0.0933	7.948	3	0.0471

Table 5

Simulation on Testing and estimation for trajectory difference with 500 runs.

Working correlation structure	Power	WASE	
		mean	sd
AR(1)	0.9460	0.6274	0.2880
Compound Symmetry	0.9180	0.9576	0.3285
Independence	0.8580	1.1518	0.5159