International Journal of Innovative Computing, Information and Control
Volume **8**, Number **8**, August **2012**
 ISSN 1349-4198
pp. 5693–5701

ASSESSING GENERALIZED LINEAR MIXED MODELS USING RESIDUAL ANALYSIS

Kuo-Chin Lin[1] and Yi-Ju Chen[2]

[1]Department of Business Administration
Tainan University of Technology
No. 529, Jhongjheng Road, YongKang Dis., Tainan City 71002, Taiwan
t20053@mail.tut.edu.tw

[2]Department of Statistics
Tamkang University
No. 151, Yingzhuan Road, New Taipei City 25137, Taiwan
ychen@stat.tku.edu.tw

Received March 2011; revised July 2011

Abstract. *A nonparametric smoothing method for assessing the adequacy of generalized linear mixed models (GLMMs) is developed. The proposed method is based on smoothing the residuals over continuous covariates to avoid the partition of continuous covariates on model checking. The global test statistic has a quadratic form and its formulae of expectation as well as variance are derived. The sampling distribution of the quadratic form test statistic is approximated by a scaled chi-squared distribution. For bandwidth selection, the leave-one-out cross-validation approach is recommendable for use. A longitudinal binary data set is utilized to demonstrate the proposed approach.*

1. **Introduction.** The analysis of longitudinal data has been remarkably flourished in the public health and social sciences for years. Longitudinal studies are employed to characterize the change in response over time and the influence of the factors on the change. The common methods for the analysis of longitudinal binary data are generalized estimating equations (GEE) models proposed by Liang and Zeger [1], and generalized linear mixed models (GLMMs). The detailed introduction of GLMMs can be referred to as McCulloch and Searle [2], Agresti [3], and Breslow and Clayton [4]. GLMMs are regarded as conditional models, whereas GEE models are treated as marginal models. Generalized linear models represent a class of fixed effects regression models for different types of response variables including continuous, dichotomous, and counts, while GLMMs are obtained from generalized linear models by incorporating random effects into the linear predictors. Some papers related tothe topic of random-effects models can be referred to as Stiratelli et al. [5], Schall [6] and Zeger and Karim [7].

It is critical for all types of regression models to assess model adequacy before making any further inferences on model parameters. Ritz [8] developed goodness-of-fit tests for mixed models based on classical goodness-of-fit statistics; see Stephens [9], and D'Agostino and Stephens [10]. In this article, we emphasize the analysis of longitudinal binary data by utilizing GLMMs, and propose a goodness-of-fit test for assessing GLMMs based on residual analysis. Many goodness-of-fit tests for GLMMs with longitudinal binary data have been vigorously developed. Vonesh et al. [11] provided a goodness-of-fit test statistic for checking the adequacy of generalized nonlinear mixed-effects models. Pan and Lin [12] considered goodness-of-fit tests of GLMMs by graphical and numerical approaches using the cumulative sums of residuals over covariates or predicted values of the response variable. Sturdivant et al. [13] presented a goodness-of-fit measure which is an unweighted sum of squares of the kernel-smoothed residuals for assessing the adequacy of GLMMs. Alonso et al. [14] proposed a family of tests to detect the misspecification in the random-effects structure of GLMMs. Huang [15] developed diagnostic tools for detecting random-effects model misspecification via coarsened data.

An alternative goodness-of-fit test based on nonparametric smoothing residuals is considered to avoid the partition of continuous covariates on model checking. The common nonparametric approaches are Nadaraya-Watson kernel smoothing, local polynomial smoothing, wavelets and splines. The pertinent papers can be referred to as Stone [16], Loader [17], Wahba [18], Daubechies [19], Green and Silverman [20], Wand and Jones [21], Fan and Gijbels [22, 23], and Simonoff [24]. Lin et al. [25] proposed a goodness-of-fit test based on nonparametric smoothing approach for GEE fitted models with continuous and categorical covariates in a longitudinal binary study.

The organization of this article proceeds as follows. The brief introduction of the current test of Pan and Lin [12] based on the cumulative sums of residuals over covariates, and a proposed goodness-of-fit test for assessing GLMMs with longitudinal binary data using nonparametric local polynomial kernel smoothing method are described in Section 2. The simulation studies are conducted to detect the sampling distribution of the proposed test statistic in Section 3. A guideline to implementation of the proposed test for longitudinal binary data is illustrated by an example, and the leave-one-out cross-validation technique for selecting a suitable bandwidth is discussed in Section 4. Lastly, conclusions and discussions are presented in Section 5.

2. **Goodness-of-Fit Test Statistics.**

2.1. **Pan-Lin's supremum test.**

Define $\{Y_{ij}\}$ as a series of binary responses for the ith subject at the jth occasion, and let $\mathbf{Y} = (\mathbf{Y}_1', \ldots, \mathbf{Y}_n')'$ with $\mathbf{Y}_i = (Y_{i1}, \ldots, Y_{in_i})'$ for $i = 1, \ldots, n$; $j = 1, \ldots, n_i$. For simplicity, we assume that $n_i = J$ for all i, and set the total sample size $N = n\,J$. The generalized linear mixed model is given by

$$\text{logit}[\text{E}(Y_{ij}|\mathbf{b}_i)] = \mathbf{X}_{ij}'\boldsymbol{\beta} + \mathbf{Z}_{ij}'\mathbf{b}_i, \tag{1}$$

where $\text{logit}(u) = \log[u/(1-u)]$. The design matrices $\mathbf{X}$ and $\mathbf{Z}$ link a $p \times 1$ vector of unknown fixed parameters $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)'$ and an $(nq) \times 1$ vector of random effects $\mathbf{b} = (\mathbf{b}_1', \ldots, \mathbf{b}_n')'$ with $\mathbf{b}_i = (b_{i1}, \ldots, b_{iq})'$ to $\mathbf{Y}$, respectively. The random effect $\mathbf{b}_i$ for subject i is assumed to follow a multivariate normal distribution, $\mathbf{b}_i \sim \mathbf{N}_q(\mathbf{0}, \mathbf{D})$. The covariance matrix $\mathbf{D}$ depends on a unknown parameter vector $\boldsymbol{\gamma}$. Let $\boldsymbol{\theta}$ represent the vector including all of the unknown parameters and denote as $\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{\gamma}')'$. Conditional on $\mathbf{b}_i$, generalized linear mixed models treat Y_{ij} as independent responses over i and j. The random effect $\mathbf{b}_i$ induces a nonnegative association among the responses for the marginal distribution average over the subjects.

Let $f_{Y|b}(Y_{ij}|\mathbf{b}_i)$ be the conditional distribution of Y_{ij} given $\mathbf{b}_i$, and $f_b(\mathbf{b}_i)$ be the density function of $\mathbf{b}_i$, respectively. The likelihood function of $\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{\gamma}')'$ is expressed as

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n}\prod_{j=1}^{n_i}\int f_{Y|b}(Y_{ij}|\mathbf{b}_i)\, f_b(\mathbf{b}_i) d\mathbf{b}_i.$$

The purpose of this article is to consider a goodness-of-fit statistic for testing H_0 : $\text{logit}[\text{E}(Y_{ij}|\mathbf{b}_i)] = \mathbf{X}_{ij}'\boldsymbol{\beta} + \mathbf{Z}_{ij}'\mathbf{b}_i$. Under H_0, the marginal means of Y_{ij} are given by $m_{ij}(\boldsymbol{\theta}) = \text{E}(Y_{ij}) = \text{E}[\text{E}(Y_{ij}|\mathbf{b}_i)]$. Based on residual analysis, Pan and Lin [12] provided a model-checking technique using cumulative sums of residuals with respect to covariates.

Define a class of stochastic process as follows:

$$W(\mathbf{x}) = n^{-1/2} \sum_{i=1}^{n} \sum_{j=1}^{n_i} I(\mathbf{X}_{ij} \leq \mathbf{x}) r_{ij}, \tag{2}$$

where $\mathbf{x}$ is a p-dimensional covariate value, $I(\mathbf{X}_{ij} \leq \mathbf{x}) = I(X_{1ij} \leq x_1, \cdots, X_{pij} \leq x_p)$, X_{lij} is the lth component of $\mathbf{X}_{ij}$, and r_{ij} is the (i,j)th component of the residual column vector $\mathbf{r}$. Here, $\mathbf{r} = \mathbf{Y} - \hat{\mathbf{m}}$ and $r_{ij} = Y_{ij} - m_{ij}(\hat{\boldsymbol{\theta}})$. Pan-Lin test is a supremum statistic, $\mathbf{T}_{PL} = \sup_{\mathbf{x}} |W(\mathbf{x})|$. Under H_0, it is obvious that the pattern of $W(\mathbf{x})$ is expected to fluctuate around zero. Large values of $\mathbf{T}_{PL}$ indicate departure from H_0. For more details, see Pan and Lin [12] and Hart [26].

2.2. **Proposed test.** Let the design matrices $\mathbf{X}$ and $\mathbf{Z}$ be partitioned into $\mathbf{X}_{N\times p} = [\mathbf{X}_1 \vdots \cdots \vdots \mathbf{X}_n]'$ and $\mathbf{Z}_{N\times(nq)} = [\mathbf{Z}_1 \bigotimes \mathbf{1}_1, \ldots, \mathbf{Z}_n \bigotimes \mathbf{1}_n]$, respectively, where the notation $\bigotimes$ is Kronecker product, $\mathbf{X}_i = (\mathbf{X}_{i1}, \ldots, \mathbf{X}_{iJ})$ with a p-dimensional covariate vector $\mathbf{X}_{ij} = (X_{ij1}, \ldots, X_{ijp})'$, $\mathbf{Z}_i = (\mathbf{Z}_{i1}, \ldots, \mathbf{Z}_{iJ})'$ with a q-dimensional covariate vector $\mathbf{Z}_{ij} = (Z_{ij1}, \ldots, Z_{ijq})'$, and $\mathbf{1}_i$ are $n \times 1$ column vectors having 1 in the ith entry and 0 elsewhere for $i = 1, \ldots, n$; $j = 1, \ldots, J$. Denote the covariance matrix of $\mathbf{Y}$ as $\mathbf{V} = \mathbf{Z}\boldsymbol{\Sigma}\mathbf{Z}' + \mathbf{U}$. The matrix $\boldsymbol{\Sigma}$ with diagonal blocks $\mathbf{D}$ is expressed as $\boldsymbol{\Sigma} = \text{diag}\{\mathbf{D}\}_1^n$, and the diagonal matrix $\mathbf{U} = \text{diag}\{\exp(u)/[1+\exp(u)]^2$, $u = \mathbf{X}'_{ij}\boldsymbol{\beta} + \mathbf{Z}'_{ij}\mathbf{b}_i$ for $i = 1, \ldots, n$; $j = 1, \ldots, J\}$. The details can be referred to Breslow and Clayton [4].

Under the regularity condition, $(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ follows an asymptotic multivariate normal distribution with zero means and a covariance matrix $\boldsymbol{\Omega}^{-1}$, where $\hat{\boldsymbol{\theta}}$ is the maximum likelihood estimator of $\boldsymbol{\theta}$, $\boldsymbol{\Omega} = \lim_{n\to\infty} \mathbf{I}(\boldsymbol{\theta})$ and $\mathbf{I}(\boldsymbol{\theta}) = -\text{E}[\partial^2 \log L(\boldsymbol{\theta})/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}']$. The information matrix can be rewritten as $\mathbf{I}(\boldsymbol{\theta}) = \text{diag}(\mathbf{I}_{11}(\boldsymbol{\theta}), \mathbf{I}_{22}(\boldsymbol{\theta}))$, where $\mathbf{I}_{11}(\boldsymbol{\theta}) = \mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$, and the kth row and lth column of $\mathbf{I}_{22}(\boldsymbol{\theta})$ is computed by $-\frac{1}{2}\,\text{tr}(\mathbf{V}^{-1}\mathbf{Z}_k\mathbf{Z}'_k\mathbf{V}^{-1}\mathbf{Z}_l\mathbf{Z}'_l)$. Zeger et al. [27] showed that the approximate marginal means for the logistic-normal case

$$m_{ij} \approx \frac{\exp(c_{ij}\mathbf{X}'_{ij}\,\boldsymbol{\beta})}{1 + \exp(c_{ij}\,\mathbf{X}'_{ij}\,\boldsymbol{\beta})}, \tag{3}$$

where $c_{ij} = |c^2\,\mathbf{D}\,\mathbf{Z}_{ij}\mathbf{Z}'_{ij} + \mathbf{I}_{nJ}|^{-q/2}$, $c = 16\sqrt{3}/(15\pi)$, and $\mathbf{I}_{nJ}$ denotes an $nJ \times nJ$ identity matrix. Analogous to (3), $m_{ij}(\hat{\boldsymbol{\theta}})$ can be approximated by

$$m_{ij}(\hat{\boldsymbol{\theta}}) \approx \exp(\hat{c}_{ij}\mathbf{X}'_{ij}\hat{\boldsymbol{\beta}}) / \left\{1 + \exp(\hat{c}_{ij}\mathbf{X}'_{ij}\hat{\boldsymbol{\beta}})\right\}, \tag{4}$$

where $\hat{c}_{ij} = |c^2\hat{\mathbf{D}}\mathbf{Z}_{ij}\mathbf{Z}'_{ij} + \mathbf{I}_{nJ}|^{-q/2}$.

The nonparametric estimators based on smoothing residuals with respect to the values of covariates $(\mathbf{x}_{11}, \ldots, \mathbf{x}_{nJ})'$ are denoted by $\hat{\mathbf{r}}_{\mathbf{h}} = (\hat{\mathbf{r}}_{\mathbf{h}}(\mathbf{x}_{11}), \ldots, \hat{\mathbf{r}}_{\mathbf{h}}(\mathbf{x}_{nJ}))'$, where $\hat{\mathbf{r}}_{\mathbf{h}}(\mathbf{x}_{ij}) = \mathbf{S}'_{\mathbf{x}_{ij}}\mathbf{r}$ for covariate values $\mathbf{x}_{ij}$, $i = 1, \cdots, n$; $j = 1, \cdots, J$. The local polynomial nonparametric smoothing matrix $\mathbf{S}'_{\mathbf{x}} = \mathbf{1}'_{dp+1}\,(\mathbf{Z}'_{\mathbf{x}}\,\mathbf{Q}_{\mathbf{x}}\,\mathbf{Z}_x)^{-1}\,\mathbf{Z}'_{\mathbf{x}}\,\mathbf{Q}_{\mathbf{x}}$, where $\mathbf{1}_{dp+1}$ is a $(d\,p{+}1) \times 1$ vector having 1 in the first entry and zero elsewhere, and $\mathbf{Z}_{\mathbf{x}}$ is an $N \times (d\,p + 1)$ matrix,

$$\mathbf{Z}_{\mathbf{x}} = \begin{bmatrix} 1 & (\mathbf{X}_{11} - \mathbf{x})' & \cdots & ((\mathbf{X}_{11} - \mathbf{x})^d)' \\ 1 & (\mathbf{X}_{12} - \mathbf{x})' & \cdots & ((\mathbf{X}_{12} - \mathbf{x})^d)' \\ \vdots & \vdots & & \vdots \\ 1 & (\mathbf{X}_{nJ} - \mathbf{x})' & \cdots & ((\mathbf{X}_{nJ} - \mathbf{x})^d)' \end{bmatrix},$$

with the degree of d. Here, $\mathbf{Q}_{\mathbf{x}} = \text{diag}\{K_{\mathbf{G}}(\mathbf{X}_{11} - \mathbf{x}), \ldots,\ K_{\mathbf{G}}(\mathbf{X}_{nJ} - \mathbf{x})\}$ and $\mathbf{G} = \text{diag}(h_1^2, \ldots, h_p^2)$. A common use of multivariate kernel function $K_{\mathbf{G}}$ is the standard p-variate normal distribution with a covariance matrix $\mathbf{G}$. The element of matrix $\mathbf{G}$ which is the nonparametric smoothing parameter can be selected by *ad hoc* or data-driven method.

We develop a new bandwidth-choice technique for the proposed test statistic based on the joint conditional log-likelihood function of Y_{ij} given b_i with leave-one-out cross-validation data-driven method in Section 4. The first-order Taylor expansion of $m_{ij}(\hat{\boldsymbol{\theta}})$ at $\boldsymbol{\theta}$ leads to

$$(Y_{ij} - m_{ij}(\hat{\boldsymbol{\theta}})) \approx (Y_{ij} - m_{ij}(\boldsymbol{\theta})) - (\partial m_{ij}(\boldsymbol{\theta})/\partial \boldsymbol{\theta})'(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}), \tag{5}$$

where $\approx$ denotes asymptotic equivalence. It can be shown that $\mathrm{E}(\mathbf{Y} - \hat{\mathbf{m}}) \approx \mathbf{0}$ and $\mathrm{Var}(\mathbf{Y} - \hat{\mathbf{m}}) \approx \mathbf{V} + \eta'\boldsymbol{\Omega}^{-1}\eta$, where $\eta = (\partial m_{11}(\boldsymbol{\theta})/\partial\boldsymbol{\theta}, \ldots, \partial m_{nJ}(\boldsymbol{\theta})/\partial\boldsymbol{\theta})$.

For any value of covariate $\mathbf{x}$, the smoothing residuals $\hat{\mathbf{r}}_{\mathbf{h}}(\mathbf{x})$ is expected to fluctuate around 0 since $\mathrm{E}(\hat{\mathbf{r}}_{\mathbf{h}}) = 0$ under H_0. The proposed goodness-of-fit test statistic with a quadratic form of $\hat{\mathbf{r}}_{\mathbf{h}}$ for GLMMs is defined by

$$\mathbf{L}_h = \frac{1}{N}\hat{\mathbf{r}}'_{\mathbf{h}}\mathrm{Cov}^{-1}(\hat{\mathbf{r}}_{\mathbf{h}})\hat{\mathbf{r}}_{\mathbf{h}}, \tag{6}$$

where $\mathrm{Cov}(\hat{\mathbf{r}}_{\mathbf{h}}) = \mathbf{S}(\mathbf{V} + \eta'\Omega^{-1}\eta)\mathbf{S}'$ and $\mathbf{S} = (\mathbf{S}_{\mathbf{x}_{11}}, \ldots, \mathbf{S}_{\mathbf{x}_{nJ}})'$. Model misspecification is revealed by the large values of $\mathbf{L}_h$. The test statistic $\mathbf{L}_h$ can be rewritten as

$$\mathbf{L}_h = \frac{1}{N}(\mathbf{Y} - \hat{\mathbf{m}})'\mathbf{K}(\mathbf{Y} - \hat{\mathbf{m}}),$$

where $\mathbf{K} = \mathbf{S}'[\mathbf{S}(\mathbf{V} + \eta'\boldsymbol{\Omega}^{-1}\eta)\mathbf{S}']^{-1}\mathbf{S}$. Based on the properties of the quadratic form $\mathbf{x}'\mathbf{K}\mathbf{x}$, it can be obtained that $\mathrm{E}(\mathbf{x}'\mathbf{K}\mathbf{x}) = \mathrm{tr}(\mathbf{K}\mathbf{V}_{\mathbf{x}}) + \mu_{\mathbf{x}}'\mathbf{K}\mu_{\mathbf{x}}$ and $\mathrm{Var}(\mathbf{x}'\mathbf{K}\mathbf{x}) = 2\mathrm{tr}(\mathbf{K}\mathbf{V}_{\mathbf{x}})^2 + 4\mu_{\mathbf{x}}'\mathbf{K}\mu_{\mathbf{x}}$, where $\mathrm{E}(\mathbf{x}) = \mu_{\mathbf{x}}$ and $\mathrm{Cov}(\mathbf{x}) = \mathbf{V}_{\mathbf{x}}$. Consequently, it can be shown that

$$\mathrm{E}(\mathbf{L}_h) = \frac{\mathrm{tr}(\mathbf{K}(\mathbf{V} + \eta'\boldsymbol{\Omega}^{-1}\eta))}{N} \tag{7}$$

and

$$\mathrm{Var}(\mathbf{L}_h) = \frac{2\mathrm{tr}(\mathbf{K}(\mathbf{V} + \eta'\boldsymbol{\Omega}^{-1}\eta))^2}{N^2}. \tag{8}$$

The asymptotic distribution of statistic $\mathbf{L}_h$ can be followed by Cox and Hinkley [28] with a scaled chi-squared distribution, $c\chi^2_\nu$, where the multiple $c = \mathrm{Var}(\mathbf{L}_h)/[2\mathrm{E}(\mathbf{L}_h)]$ and the degrees of freedom $\nu = 2\mathrm{E}^2(\mathbf{L}_h)/\mathrm{Var}(\mathbf{L}_h)$.

For continuous and categorical covariates in GLMMs, we stratify data by categorical variables mentioned by le Cessie and van Houwelingen [29] if the number of categories is small relative to the size of observations. A goodness-of-fit statistic for testing GLMMs with continuous and categorical covariates is given by $\mathbf{L}^\star_h = \sum_g \mathbf{L}_{h,g}$, where $\mathbf{L}_{h,g}$ is the test statistic using smoothing the residuals of the fitted model in the gth category. It can be straightforwardly obtained the expectation and variance of $\mathbf{L}^\star_h$, and the sampling distribution of $\mathbf{L}^\star_h$ can be approximated by a scaled chi-squared distribution.

3. **Simulation Study.** Two simulation studies are conducted to detect the behavior of asymptotic sampling distribution of the proposed statistic $\mathbf{L}^\star_h$, and to compare the performance between Pan-Lin's supremum test and the proposed test, where two simulated longitudinal binary data sets are generated from the following null models:

$$\text{Model 1: } \mathrm{logit}[\mathrm{E}(Y_{ij}|b_i)] = 1.5 - 0.5X_{1i} + X_{2ij} + b_i,$$
$$\text{Model 2: } \mathrm{logit}[\mathrm{E}(Y_{ij}|b_i)] = -0.5X_{1i} + 0.5X_{2ij} + b_i.$$

In Model 1, X_{1i} is a time-independent covariate and follows a Bernoulli distribution with probability of success 0.4. $\mathbf{X}_{2i} = (X_{2i1}, X_{2i2}, X_{2i3})'$ is a time-dependent covariate vector. The conditional distribution of $\mathbf{X}_{2i}$ given X_{1i} has a multivariate normal distribution with a mean vector $(X_{1i}, X_{1i}, X_{1i})'$ and a correlation matrix with the pairwise correlations of 0.5. The random effect b_i follows the standard normal distribution for $i = 1, \cdots, n$. In Model 2, X_{1i} and X_{2ij} are two continuous covariates, and follow a uniform distribution on

$(-2,2)$ and $b_i \sim N(0,1)$ for $i = 1, \cdots, n$; $j = 1, 2$, respectively. The correlations between two occasions are assumed to be 0.5. Each setting is performed for 1000 replications.

The proportions of rejection for Pan-Lin's supremum test and the proposed test compared with the significance levels of a scaled chi-squared distribution are summarized in Tables 1 and 2. They indicate that most of empirical rejection levels have good approximations of significance levels for all bandwidths, and the supremum test has the empirical sizes near the nominal levels. The slight differences may be due to simulation error. Note that it needs to consider the paired bandwidths for nonparametric smoothing with respect to the continuous covariates in Model 2.

Table 1. The proportions of rejection for the supremum test $\mathbf{T}_{PL}$ and the proposed test $\mathbf{L}_h^\star$ in Model 1 for various sample sizes, bandwidths and significance levels

		$\mathbf{T}_{PL}$		$\mathbf{L}_h^\star$		
n	α			Bandwidth h		
			1.0	1.5	2.0	2.5
50	0.010	0.009	0.010	0.017	0.014	0.017
	0.025	0.023	0.025	0.029	0.032	0.029
	0.050	0.046	0.053	0.054	0.053	0.054
	0.100	0.102	0.093	0.089	0.091	0.089
100	0.010	0.013	0.014	0.014	0.011	0.014
	0.025	0.026	0.024	0.027	0.027	0.027
	0.050	0.048	0.042	0.045	0.046	0.045
	0.100	0.101	0.081	0.099	0.096	0.099
200	0.010	0.011	0.012	0.014	0.011	0.014
	0.025	0.026	0.022	0.025	0.028	0.025
	0.050	0.054	0.053	0.049	0.052	0.049
	0.100	0.097	0.102	0.098	0.106	0.098

Table 2. The proportions of rejection for the supremum test $\mathbf{T}_{PL}$ and the proposed test $\mathbf{L}_h^\star$ in Model 2 for various sample sizes, bandwidths and significance levels

		$\mathbf{T}_{PL}$		$\mathbf{L}_h^\star$		
n	α			Bandwidth (h,h)		
			1.0	1.5	2.0	2.5
50	0.010	0.011	0.011	0.008	0.009	0.009
	0.025	0.026	0.027	0.027	0.027	0.021
	0.050	0.050	0.048	0.049	0.046	0.043
	0.100	0.103	0.100	0.101	0.093	0.095
100	0.010	0.011	0.015	0.013	0.008	0.007
	0.025	0.024	0.027	0.025	0.022	0.020
	0.050	0.049	0.058	0.045	0.049	0.055
	0.100	0.103	0.094	0.102	0.106	0.107
200	0.010	0.009	0.012	0.011	0.010	0.009
	0.025	0.025	0.027	0.031	0.023	0.025
	0.050	0.052	0.040	0.041	0.040	0.042
	0.100	0.096	0.082	0.078	0.081	0.086

4. **An Example.** A longitudinal binary data set, analyzed by Koch et al. [30], Davis [31] and also cited by Lin et al. [25], is used to demonstrate the application of the proposed goodness-of-fit test for GLMMs using nonparametric smoothing residuals. A total of 111 patients from two different centers were randomly assigned to receive either placebo (57 patients) or an active treatment (54 patients), and the patient's respiratory status was examined (1 = good, 0 = poor) at baseline and four visits during treatment. The GLMM for respiratory illness data,

$$\text{logit}[\text{E}(Y_{ij}|b_i)] = \beta_0 + \beta_1\text{C}_j + \beta_2\text{T}_j + \beta_3\text{G}_j + \beta_4\text{B}_j + \beta_5\text{A}_j + b_i,$$

includes four binary time-independent covariates: center ($C = 0$ for center 1, $C = 1$ for center 2), treatment ($T = 0$ for placebo group, $T = 1$ for active treatment), gender ($G = 0$ for female, $G = 1$ for male) and baseline ($B = 0$ for response 'poor', $B = 1$ for response 'good'), and a continuous time-independent covariate A, the age of the patient (in years at baseline) for $i = 1, \cdots, 111$; $j = 1, \cdots, 4$. Denote the binary repeated patient's respiratory status outcomes at four visits during treatment as $\mathbf{Y}_i = (Y_{i1}, \cdots, Y_{i4})'$. The estimates and confidence intervals of odds ratio for model parameters, the corresponding standard errors, and the estimated standard deviation of the random intercept as well as the p-value of the proposed test are shown in Table 3.

Table 3. Estimates, standard errors and conference intervals for GLMM model parameters as well as the results of the proposed test with an optimal bandwidth $h_{opt} = 0.49$

Covariate	$\hat{\theta}$	S.E.	95% CI for e^{β}
Intercept	–0.754	0.448	(0.196, 1.132)
Center	0.694	0.244	(1.241, 3.230)
Treatment	1.308	0.241	(2.306, 5.932)
Gender	–0.127	0.299	(0.490, 1.583)
Baseline	1.909	0.246	(4.166, 10.926)
Age	–0.020	0.009	(0.963, 0.998)
σ	0.470	0.118	–
Test statistic	Mean	Variance	p-value
0.328	3.457	1.609	0.999

The selection of an optimal smoothing parameter is a crucial issue for smoothing methods due to its influence on the performance of the test. The commonly used technique for selecting an optimal bandwidth, based on maximizing the cross-validation function with respect to h, is the leave-one-out cross-validation data-driven method, referred to Fan and Gijbels [22], and Jones et al. [32]. Let $\hat{m}_{ij}^{(-i)}$ represent the estimate of the marginal means of Y_{ij}, $m_{ij}(\boldsymbol{\theta}) = \text{E}(Y_{ij}) = \text{E}[\text{E}(Y_{ij}|\mathbf{b}_i)]$, without considering the ith patient. The optimal bandwidth h_{opt} is obtained by maximizing the log-likelihood function of Y_{ij}, $CV(h)$, where

$$CV(h) = \sum_{i=1}^{n}\sum_{j=1}^{n_i}\left[Y_{ij}\log\left(\hat{m}_{ij}^{(-i)}\right) + (1 - Y_{ij})\log\left(\left(1 - \hat{m}_{ij}^{(-i)}\right)\right)\right]. \quad (9)$$

Here $\hat{m}_{ij}^{(-i)}$ is computed by the nonparametric smoother of $\text{E}[\exp(u_{ij})/(1 + \exp(u_{ij}))|\mathbf{X}_{ij}]$ with $u_{ij} = \mathbf{X}_{ij}'\hat{\boldsymbol{\beta}} + b_i$ and $b_i \sim N(0, \hat{\sigma}^2)$. Figure 1 displays the log-likelihood function curve $CV(h)$ versus bandwidth h based on kernel and local linear smoothing, respectively. The optimal bandwidth h_{opt} relies on model specification, and its kernel smoothing, $h_{opt} = 0.49$.

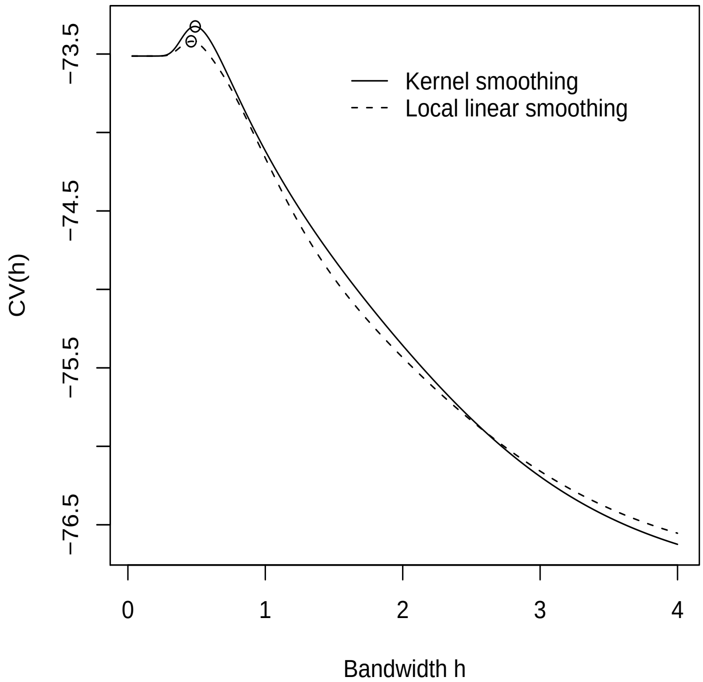

FIGURE 1. Plot of the log-likelihood function based on leave-one-out cross-validation versus bandwidth h. The optimal bandwidths are 0.49 and 0.46 with kernel and local linear smoothing, respectively.

The proposed test statistic $\mathbf{L}_h^\star = 0.328$ with p-value $= 0.999$. It concludes that the model is decent.

The results in Table 3 reveal that the respiratory status is significantly affected by all of the covariates. The center effect is inferred by the estimated odds ratio $\exp(0.694) = 2.002$ and its 95% confidence interval $(1.241, 3.230)$. The treatment effect is inferred by the estimated odds ratio $\exp(1.308) = 3.699$ and its 95% confidence interval $(2.306, 5.932)$. Likewise, the inferences of the other covariates can be obtained in the same manner. The results also imply that the younger female patient with response 'good' at baseline has improved respiratory status after receiving active treatment in center 2. The estimated standard deviation of random effect is 0.470 (p-value ≈ 0), and it shows that the degree of heterogeneity among patients with good or poor respiratory status is manifest.

5. **Conclusion and Discussion.** In contrast with the method of Lin et al. [25] using GEE approach, this article provides an alternative model-checking technique for generalized linear mixed models based on nonparametric local polynomial smoothing residuals. The proposed test can be regarded as a complement to Pan-Lin's [12] supremum test rather than a competitor. The proposed statistic is easy to compute, and is applicable for checking the fitted models with continuous and categorical covariates. This global measure goodness-of-fit test statistic avoids partitioning the covariate space. The proportions of rejection for the proposed test statistic can be well approximated by its asymptotic scaled chi-squared distribution.

In addition, a real longitudinal binary data set is employed to demonstrate the proposed goodness-of-fit testing procedure. The extent of bandwidth h depends on smoothing mean function, and the selection of smoothing parameter for nonparametric smoothing methods is crucial. A bandwidth-choice technique with leave-one-out cross-validation method is used for the proposed test based on the joint conditional log-likelihood function of Y_{ij} given b_i.

The evaluation of the random effect may be of interest, which is not the focus of this article. Verbeke and Lesaffre [33, 34] obtained the maximum likelihood estimators for random effects, and investigated the effect on misspecification of the random-effects distribution in linear mixed models with longitudinal data. Litière et al. [35, 36] and Huang [15] developed tests for detecting the random-effects misspecification in GLMMs, and Alonso et al. [14] provided a family of tests to detect misspecification in the random-effects structure of generalized linear mixed models. The proposed method may be extended to longitudinal ordinal data. Regarding the analysis of longitudinal ordinal data, it can be referred to Chen et al. [37], Lin et al. [38] and Tuan et al. [39]

Acknowledgments. The authors are grateful to the reviewers for their valuable comments. We also thank Professor Junzo Watada and Professor Yan Shi for their helpful comments and kind assistance. This work was partially supported by National Science Council Grant NSC-98-2118-M-165-001.

REFERENCES

[1] K. Y. Liang and S. L. Zeger, Longitudinal data analysis using generalized linear models, *Biometrika*, vol.73, pp.13-22, 1986.
[2] C. E. McCulloch, S. R. Searle and J. M. Neuhaus, *Generalized, Linear, and Mixed Models*, Wiley, New York, 2008.
[3] A. Agresti, *Categorical Data Analysis*, Wiley, New York, 2002.
[4] N. E. Breslow and D. G. Clayton, Approximate inference in generalized linear mixed models, *Journal of the American Statistical Association*, vol.88, pp.9-25, 1993.
[5] R. Stiratelli, N. Laird and J. H. Ware, Random effects models for serial observations with binary responses, *Biometrics*, vol.40, pp.961-971, 1984.
[6] R. Schall, Estimation in generalized linear models with random effects, *Biometrika*, vol.78, pp.719-727, 1991.
[7] S. L. Zeger and M. R. Karim, Generalized linear models with random effects: A Gibbs sampling approach, *Journal of the American Statistical Association*, vol.86, pp.79-86, 1991.
[8] C. Ritz, Goodness-of-fit tests for mixed models, *The Scandinavian Journal of Statistics*, vol.31, pp.443-458, 2004.
[9] M. A. Stephens, EDF statistics for goodness of fit and some comparison, *Journal of the American Statistical Association*, vol.69, pp.730-737, 1974.
[10] R. D'Agostino and M. A. Stephens, *Goodness-of-Fit Techniques*, Marcel Dekker, New York, 1986.
[11] E. F. Vonesh, V. M. Chinchilli and K. Pu, Goodness-of-fit in generalized nonlinear mixed-effects models, *Biometrics*, vol.52, pp.572-587, 1996.
[12] Z. Pan and D. Y. Lin, Goodness-of-fit methods for generalized linear mixed models, *Biometrics*, vol.61, pp.1000-1009, 2005.
[13] R. X. Sturdivant, J. J. Rotella and R. E. Russell, A smoothed residual based goodness-of-fit statistic for nest-survival models, *Studies in Avian Biology*, vol.34, pp.45-54, 2006.
[14] A. Alonso, S. Litière and G. Molenberghs, A family of tests to detect misspecifications in the random-effects structure of generalized linear mixed models, *Computational Statistics & Data Analysis*, vol.52, pp.4474-4486, 2008.
[15] X. Z. Huang, Detecting random-effects model misspecification via coarsened data, *Computational Statistics & Data Analysis*, vol.55, pp.703-714, 2011.
[16] C. J. Stone, Consistent nonparametric regression, *The Annual of Statistics*, vol.5, pp.595-620, 1977.
[17] C. Loader, *Local Regression and Likelihood*, Springer-Verlag, New York, 1999.
[18] G. Wahba, *Spline Models for Observational Data*, SIAM, Philadelphia, 1990.
[19] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Philadelphia, 1992.
[20] P. J. Green and B. W. Silverman, *Nonparametric Regression and Generalized Linear Models*, Chapman & Hall, London, 1994.

[21] M. P. Wand and M. C. Jones, *Kernel Smoothing*, Chapman & Hall, London, 1995.

[22] J. Fan and I. Gijbels, Data-driven bandwidth selection in local polynomial fitting: Variable bandwidth and spatial adaptation, *Journal of the Royal Statistical Society, Series B*, vol.57, pp.371-394, 1995.

[23] J. Fan and I. Gijbels, *Local Polynomial Modelling and Its Applications*, Chapman & Hall, London, 1996.

[24] J. S. Simonoff, *Smoothing Methods in Statistics*, Springer-Verlag, New York, 1996.

[25] K. C. Lin, Y. J. Chen and Y. Shyr, A nonparametric smoothing method for assessing GEE models with longitudinal binary data, *Statistics in Medicine*, vol.27, pp.4428-4439, 2008.

[26] J. D. Hart, *Nonparametric Smoothing and Lack-of-Fit Tests*, Springer-Verlag, New York, 1997.

[27] S. L. Zeger, K. Y. Liang and P. S. Albert, Models for longitudinal data: A generalized estimating equation approach, *Biometrics*, vol.44, pp.1049-1060, 1988.

[28] D. R. Cox and D. V. Hinkley, *Theoretical Statistics*, Chapman & Hall, London, 1974.

[29] S. le Cessie and J. C. van Houwelingen, A goodness-of-fit test for binary regression models based on smoothing methods, *Biometrics*, vol.47, pp.1267-1282, 1991.

[30] G. G. Koch, G. J. Carr, I. A. Amara, M. E. Stokes and T. J. Uryniak, Categorical data analysis, in *Statistical Methodology in the Pharmaceutical Sciences*, D. A. Berry (ed.), New York, Marcel Dekker, 1990.

[31] C. S. Davis, Semi-parametric and non-parametric methods for the analysis of repeated measurements with applications to clinical trials, *Statistics in Medicine*, vol.10, pp.1959-1980, 1991.

[32] M. C. Jones, J. S. Marron and S. J. Sheather, A brief survey of bandwidth selection for density estimation, *Journal of the American Statistical Association*, vol.91, pp.401-407, 1996.

[33] G. Verbeke and E. Lesaffre, A linear mixed-effects model with heterogeneity in the random-effects population, *Journal of the American Statistical Association*, vol.91, pp.217-221, 1996.

[34] G. Verbeke and E. Lesaffre, The effect of misspecifying the random-effects distribution in linear mixed models for longitudinal data, *Computational Statistics & Data Analysis*, vol.23, pp.541-556, 1997.

[35] S. Litière, A. Alonso and G. Molenberghs, Type I and type II error under random-effects misspecification in generalized linear mixed models, *Biometrics*, vol.63, pp.1038-1044, 2007.

[36] S. Litière, A. Alonso and G. Molenberghs, The impact of a misspecification random-effects distribution on the estimation and the performance of inferential procedures in generalized linear mixed models, *Statistics in Medicine*, vol.27, pp.3125-3144, 2008.

[37] Y. J. Chen, K. C. Lin and J. J. Lin, Group sequential analysis of incomplete longitudinal ordinal data, *ICIC Express Letters*, vol.3, no.4(B), pp.1453-1458, 2009.

[38] K. C. Lin, Y. J. Chen and C. Y. Lui, Goodness-of-fit tests for GEE models with longitudinal ordinal data using a global odds ratio, *ICIC Express Letters*, vol.5, no.6, pp.1821-1826, 2011.

[39] L. W. Tuan, Y. J. Chen, P. L. Li and K. C. Lin, Evaluation of multiple imputation for longitudinal ordinal data under MCAR and MAR missing-data mechanisms, *ICIC Express Letters*, vol.5, no.6, pp.1833-1838, 2011.